[File 2] INSPEC 1898-2008/Jun W2

[File 6] NTIS 1964-2008/Jul W3

[File 8] Ei Compendex(R) 1884-2008/Jun W5

[File 34] SciSearch(R) Cited Ref Sci 1990-2008/Jun W5

[File 35] Dissertation Abs Online 1861-2008/Nov

[File 56] Computer and Information Systems Abstracts 1966-2008/Jun

[File 57] Electronics & Communications Abstracts 1966-2008/May

[File 65] Inside Conferences 1993-2008/Jul 09

[File 95] TEME-Technology & Management 1989-2008/Jun W5

[File 98] General Sci Abs 1984-2008/May

[File 553] Wilson Bus. Abs. 1982-2008/May

[File 99] Wilson Appl. Sci & Tech Abs 1983-2008/Apr

[File 144] Pascal 1973-2008/Jul W1

[File 239] Mathsci 1940-2008/Jul

[File 256] TecInfoSource 82-2008/Oct

[File 434] SciSearch(R) Cited Ref Sci 1974-1989/Dec

[File 583] Gale Group Globalbase(TM) 1986-2002/Dec 13

**File 583: This file is no longer updating as of 12-13-2002.*

[File 603] Newspaper Abstracts 1984-1988

**File 603: This is a closed file.*

[File 483] Newspaper Abs Daily 1986-2008/Jul 09


```
Set      Items   Description
S1      333103   S MIC OR MICS OR TRANSDUCER?? OR MICROPHONE??
S2        7323   S MEL()FREQUENCY()CEPSTRAL()COEFFICIENT?? OR MFCC OR CEPTSTRUM
OR CEPSTRAL
S3       63333   S (WEIGHT??? OR SCORE?? OR SCORING OR NUMERIC? OR SUM?? OR
ADD?? OR SCALAR?? OR COEFFICIENT??)(5N)VECTOR??
```

```
S4   10690117   S STATISTIC? OR STOCHASTIC? OR PROBABILITY OR LIKELY OR
LIKELIHOOD OR PROXIM? OR APPROXIMAT??? OR ESTIMAT???
S5     172615   S AU=(LIU, Z? OR LIU Z? OR SINCLAIR, M? OR SINCLAIR M? OR
ACERO, A? OR ACERO A? OR HUANG, X? OR HUANG X? OR DROPPO, J? OR DROPPO J? OR
DENG, L? OR DENG L? OR ZHANG, Z? OR ZHANG Z? OR ZHENG, Y? OR ZHENG Y?)
S6     177519   S (REDUCE?? OR MIN OR MINIMUM OR MINIMIS?? OR MINIMIZ??? OR
REDUCTION OR REDUCING OR FILTER??? OR ELIMINAT??? OR ATTENUAT???)(3N)(NOISE??
OR ARTIFACT??)
S7        262   S S1 AND S2
S8         10   S S7 AND S3
S9          6   S S8 AND S4
S10         3   RD  (unique items)
S11         1   S S10 NOT PY>2003
S12         4   S S8 NOT (S9 OR PY>2003)
S13         4   RD  (unique items)
S14       366   S S3 AND S4 AND S6
S15         9   S S14 AND S1
S16         3   S S15 NOT (S9 OR S13 OR PY>2003)
S17         2   RD  (unique items)
S18         3   S S14 AND S2
S19         2   S S18 NOT (S9 OR S13 OR S17 OR PY>2003)
S20         7   S S5 AND S14
S21         0   S S20 NOT (S9 OR S13 OR S17 OR S19 OR PY>2003)
S22        12   S S5 AND S7
S23        10   RD  (unique items)
S24         7   S S23 NOT (S9 OR S13 OR S17 OR S19 OR PY>2003)
```

11/3,K/1 (Item 1 from file: 2) Links
Fulltext available through: STIC Full Text Retrieval Options
INSPEC

05394481 INSPEC Abstract Number: B9306-6130-014, C9306-1250C-005
Title: Filterbank-energy estimation using mixture and Markov models for recognition of noisy speech
Author Erell, A.; Weintraub, M.
Author Affiliation: SRI Inst., Menlo Park, CA, USA
Journal: IEEE Transactions on Speech and Audio Processing vol.1, no.1 p. 68-76

Publication Date: Jan. 1993 Country of Publication: USA
CODEN: IESPEJ ISSN: 1063-6676
U.S. Copyright Clearance Center Code: 1063-6676/93/$03.00
Language: English
Subfile: B C
Title: Filterbank-energy estimation using mixture and Markov models for recognition of noisy speech
Abstract: An estimation algorithm for noise robust speech recognition, the minimum mean log spectral distance (MMLSD), is presented. The estimation is matched to the recognizer by seeking to minimize the average distortion as measured by a Euclidean distance between filterbank log-energy vectors, approximating the weighted-cepstral distance used by the recognizer. The estimation is computed using a clean speech spectral probability distribution, estimated from a database, and a stationary, ARMA model for the noise. When trained on clean speech and tested with additive white noise at 10-dB... ...the recognizer at the same constant 10-dB SNR. The algorithm is also highly efficient with a quasi-stationary environmental noise, recorded with a desktop microphone, and requires almost no tuning to differences between this noise and the computer-generated white noise.
Identifiers: ...filterbank energy estimation; ... ...estimation algorithm... ...weighted-cepstral distance... ...speech spectral probability distribution... ...desktop microphone;
Astronomical Objects:

13/3,K/1 (Item 1 from file: 2) Links
INSPEC

09131590 INSPEC Abstract Number: B2004-11-6130E-021, C2004-11-5260S-020

Title: Automatic identification of a speaker from the record of his speech in the presence of passive and active noises
Author Zinchenko, E.Y.; Popov, A.M.
Author Affiliation: Dept. of Comput. Math. & Cybern., Moscow Univ., Russia
Journal: Vestnik Moskovskogo Universiteta, Seriya 15 (Vychislitel'naya Matematika i Kibernetika) no.2
Publisher: Allerton Press ,
Publication Date: 2003 Country of Publication: Russia
CODEN: VMUKD8 ISSN: 0201-7385
Material Identity Number: M248-2004-001
Translated in: Moscow University Computational Mathematics and Cybernetics no.2 p. 34-41
Publication Date: 2003 Country of Publication: USA
CODEN: MUCTD4 ISSN: 0278-6419
SICI of Translation: 0278-6419(2003)2L.34:AISF;1-2
U.S. Copyright Clearance Center Code: 0278-6419/03/$50.00
Language: English
Subfile: B C

Abstract: ...of identification of a speaker under the conditions of different kinds of noise. The approach is based on the use of quantization of the feature vectors in the space of cepstral coefficients with a subsequent clustering of the data. The chosen system of identification shows a hundred percent recognition of a speaker in the well-chosen laboratory... ...however, that the exactness of identification is lost in the presence of different kinds of noise. Natural noise are considered which deteriorate the quality of microphone recording as well as active noises which are due to the presence of other speakers when the recording is carried out.
Descriptors: ...cepstral analysis
Identifiers: ...cepstral coefficients... ...microphone recording

13/3,K/2 (Item 2 from file: 2) Links
INSPEC

05917055 INSPEC Abstract Number: B9505-6130-182, C9505-1250C-086
Title: Environment normalization for robust speech recognition using direct cepstral comparison
Author Fu-Hua Liu; Stern, R.M.; Acero, A.; Moreno, P.J.
Author Affiliation: Dept. of Electr. & Comput. Eng., Carnegie Mellon Univ., Pittsburgh, PA, USA
Part vol.2 p. II/61-4 vol.2
Publisher: IEEE , New York, NY, USA
Publication Date: 1994 Country of Publication: USA 6 vol. 3382 pp.
ISBN: 0 7803 1775 0
U.S. Copyright Clearance Center Code: 0 7803 1775 0/94/$3.00
Conference Title: Proceedings of ICASSP '94. IEEE International Conference on Acoustics, Speech and Signal Processing
Conference Sponsor: IEEE Signal Process. Soc
Conference Date: 19-22 April 1994 Conference Location: Adelaide, SA, Australia
Language: English
Subfile: B C

Title: Environment normalization for robust speech recognition using direct cepstral comparison
Abstract: ...describe and evaluate a series of new algorithms that compensate for the effects of unknown acoustical environments or changes in environment. The algorithms use compensation vectors that are added to the cepstral representations of speech that is input to a speech recognition system. While these vectors are computed from direct frame-by-frame comparisons of cepstra of... ...phonetic identity. The compensation algorithms are evaluated using the 1992 ARPA 5000 word WSJ/CSR corpus. The best system combines phoneme-based and SNR-based cepstral compensation with cepstral mean normalization, and provides a 66.8% reduction in error rate over baseline processing when tested using a standard suite of unknown microphones.
Descriptors: cepstral analysis...
Identifiers: ...direct cepstral comparison... ...cepstral mean normalization

13/3,K/3 (Item 1 from file: 8) Links
Fulltext available through: STIC Full Text Retrieval Options
Ei Compendex(R)

07102366 E.I. No: EIP95032605354

Title: Environment normalization for robust speech recognition using direct cepstral comparison

Author: Liu, Fu-Hua; Stern, Richard M.; Acero, Alejandro; Moreno, Pedro J.
Corporate Source: Carnegie Mellon Univ, Pittsburgh, PA, USA
Conference Title: Proceedings of the 1994 IEEE International Conference on Acoustics, Speech and Signal Processing. Part 2 (of 6)
Conference Location: Adelaide, Aust Conference Date: 19940419-19940422
E.I. Conference No.: 42612
Source: Proceedings - ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing v 2 1994. IEEE, Piscataway, NJ, USA,94CH3387-8. p 61-64
Publication Year: 1994
CODEN: IPRODJ ISSN: 0736-7791
Language: English
Title: Environment normalization for robust speech recognition using direct cepstral comparison

Abstract: ...describe and evaluate a series of new algorithms that compensate for the effects of unknown acoustical environments or changes in environment. The algorithms use compensation vectors that are added to the cepstral representations of speech that is input to a speech recognition system. While these vectors are computed from direct frame-by-frame comparisons of cepstra of... ...presumed phonetic identity. The compensation algorithms are evaluated using the 1992 ARPA 5000-word WSJ/CSR corpus. The best system combines phonemebased and SNR-based cepstral compensation with cepstral mean normalization, and provides a 66.8% reduction in error rate over baseline processing when tested using a standard suite of unknown microphones. (Author abstract) 8 Refs.
Descriptors:
Identifiers: Acoustical environment; Cepstral representation of speech; Presumed phonetic identity; Cepstral normalization algorithm; Codeword; Linear filtering; Additive noise; Cepstral coefficient; Cepstral vectors; Additive correction
Identifiers:

13/3,K/4 (Item 1 from file: 144) Links
Pascal

11484063 PASCAL No.: 94-0322058

Study of phonetic classification/recognition performance on data collected with different microphones

CHANG Jane; ZUE Victor
Spoken Language Systems Group, Lab. for Comput. Sci., MIT, 545 Technology Sq., Cambridge, MA 02139
The 127th Meeting of the Acoustical Society of America (Cambridge, Massachusetts (USA))
1994-06-06/1994-06-10
Journal: Journal of the Acoustical Society of America, 1994-05 , 95 (5) 2877-2877
Language: English

Study of phonetic classification/recognition performance on data collected with different microphones

The goal of this study is to seek an understanding of the effects of microphone variations on the MIT segment-based speech recognition system, summit. Specifically, phonetic classification and recognition performance are evaluated on utterances extracted from the

timit corpus. The timit
corpus offers phonetically-transcribed and time-aligned data for three different microphones
-a Sennheiser close-talking, noise-canceling microphone , a Bruel and Kajar (B&K)
far-field pressure microphone, and a telephone handset (plus channel distortion). These
transducers cause different convolutional, additive, and bandwidth effects in the speech
waveform. Experimental procedures are established to measure and analyze system performance under
variable training and testing conditions. Classification uses Gaussian models on a feature
vector consisting of Mel-frequency cepstral coefficients and
their time derivatives, plus duration. The experiments show that performance in phonetic
classification and recognition degrades from the Sennheiser (27% classification error) to the...

English Descriptors: Experimental study; Speech recognition; Microphones; Variations;
Performance testing; Classification

French Descriptors: Etude experimentale; 4372N; 4338K; Reconnaissance parole; Microphone;
Variation; Essai performance; Classification

17/3,K/1 (Item 1 from file: 8) Links
Ei Compendex(R)

08915632 E.I. No: EIP01436696408
Title: Blind source separation combining frequency-domain ICA and beamforming

Author: Saruwatari, H.; Kurita, S.; Takeda, K.
Corporate Source: Grad. School of Information Science Nara Inst. of Science and Technology, Ikoma-shi, Nara 630-0101, Japan
Conference Title: 2001 IEEE Interntional Conference on Acoustics, Speech, and Signal Processing
Conference Location: Salt Lake, UT, United States Conference Date: 20010507-20010511
E.I. Conference No.: 58545
Source: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings v 5 2001. p 2733-2736 (IEEE cat n 01CH37221)
Publication Year: 2001
CODEN: IPRODJ ISSN: 0736-7791
Language: English
Abstract: In this paper, we describe a new method of blind source separation (BSS) on a microphone array combining subband independent component analysis (ICA) and beamforming. The proposed array system consists of the following three sections: (1) subband-ICA-based BSS section with direction-of-arrival (DOA) estimation, (2) null beamforming section based on the estimated DOA information, and (3) integration of (1) and (2) based on the algorithm diversity. Using this technique, we can resolve the low-convergence problem through optimization in ICA. The results of the signal separation experiments reveal that the noise reduction rate (NRR) of about 18 dB is obtained under the nonreverberant condition, and NRRs of 8 dB and 6 dB are obtained in the case...
Descriptors: *Blind source separation; Independent component analysis; Frequency domain analysis; Microphones; Acoustic arrays; Algorithms; Convergence of numerical methods; Optimization; Acoustic noise; Reverberation; Vectors
Identifiers: Microphone array; Beamforming; Noise reduction rate; Direction-of-arrival
Identifiers:

17/3,K/2 (Item 1 from file: 95) Links
Fulltext available through: STIC Full Text Retrieval Options
TEME-Technology & Management

01736446 20030310652
The performance surface in filtered nonlinear mean-square estimation

Costa, MH; Bermudez, JCM; Bershad, NJ
Grupo de Engenharia Biomedica, Univ. Catolica de Pelotas, Brazil
IEEE Transactions on Circuits and Systems I, Fundamental Theory and Applications, v50, n3, pp445-447 , 2003
Document type: journal article Language: English
Record type: Abstract
ISSN: 1057-7122
The performance surface in filtered nonlinear mean-square estimation

Abstract:
This brief investigates the properties of the performance surface for the problem of linearly constrained nonlinear mean-square estimation of a random sequence. The problem studied has direct application to the study of active noise control systems when the transducers are driven into nonlinear behavior. A deterministic expression is derived for the mean-square error (MSE) surface as a function of the nonlinearity parameter for Gaussian inputs. It is demonstrated that the surface is unimodal, and expressions are determined for the optimum weight vector and for the minimum MSE.

Descriptors: ADAPTIVE FILTERS; ADAPTIVE SIGNAL PROCESSING; FILTER THEORY; RANDOM SEQUENCE; ACOUSTIC NOISE REDUCTION
Identifiers:

19/3,K/1 (Item 1 from file: 2) Links
Fulltext available through: STIC Full Text Retrieval Options
INSPEC

08296313 INSPEC Abstract Number: B2002-07-6130E-050, C2002-07-1250C-027
Title: An efficient algorithm for automatic robust speech recognition
Author Kotnik, B.; Kacic, Z.; Horvat, B.
Author Affiliation: Fakulteta za elektrotehniko, Maribor Univ., Slovenia
Journal: Elektrotehniski Vestnik vol.69, no.1 p. 69-74
Publisher: Electrotech. Soc. Slovenia ,
Publication Date: 2002 Country of Publication: Slovenia
CODEN: ELVEA2 ISSN: 0013-5852
SICI: 0013-5852(2002)69:1L.69:EAAR;1-N
Material Identity Number: E040-2002-003
Language: Slovenian
Subfile: B C

Abstract: ...an automatic speech recogniser that works well in a wide range of unexpected or adverse environments. In this paper, we present an effective two-stage noise reduction procedure, which uses time and spectral domain processing and achieves a trade-off between effective noise reduction and low computational load for real-time operations. At the first stage, a novel weighting function is used to reduce the effect of additive noise on speech in time domain. At the second stage, a spectral subtraction method based on minimum statistics is used. The last step of the proposed algorithm is a mel cepstrum feature extraction procedure. The feature vector consists of 12 mel cepstrum coefficients and the energy parameter. The Slovenian fixed telephone database (FDB) SpeechDat II, the German SpeechDat II FDB as well as Spanish SpeechDat 11 FDB were...
Descriptors: ...cepstral analysis
Identifiers: ...two-stage noise reduction procedure... ...noise reduction; ... ...minimum statistics;

19/3,K/2 (Item 2 from file: 2) Links
INSPEC

08048354 INSPEC Abstract Number: B2001-11-6130C-004
Title: Speech coding and noise reduction using ICA-based speech features

Author Jong-Hwan Lee; Ho-Young Jung; Te-Won Lee; Soo-Young Lee
Author Affiliation: Dept. of Electr. Eng., Korea Adv. Inst. of Sci. & Technol., Taejon, South Korea
Conference Title: Second International Workshop on Independent Component Analysis and Blind Signal Separation. Proceedings p. 417-21
Editor(s): Pajunen, P.; Karhunen, J.
Publisher: Helsinki Univ. Technol , Espoo, Finland
Publication Date: 2000 Country of Publication: Finland 647 pp.
ISBN: 951 22 5017 9 Material Identity Number: XX-2000-01026
Conference Title: Proceedings of International Workshop on Independent Component Analysis and Blind Signal Separation (ICA 2000)
Conference Date: 19-22 June 2000 Conference Location: Helsinki, Finland
Language: English
Subfile: B
Copyright 2001, IEE
Title: Speech coding and noise reduction using ICA-based speech features
Abstract: ...When independent component analysis is applied to speech signals for efficient encoding the adapted basis vectors have Gabor-like features. Then only a few active coefficients of the trained basis vectors are sufficient for encoding the speech signals. Those trained speech features can he used in automatic speech recognition systems, and the proposed method gives better recognition rates than conventional mel-frequency cepstral coefficients (MFCC) features. Trained basis vectors can be also applied for the removal of Gaussian noise. Speech signals corrupted by additive white Gaussian noise show much improvement in the signal-to-noise ratio (SNR) after the denoising process. Then, these denoised speech features show better recognition performance than MFCC features.
Descriptors: ...statistical analysis
Identifiers: ...noise reduction;

24/3,K/1 (Item 1 from file: 2) Links
INSPEC
(c) 2008 Institution of Electrical Engineers. All rights reserved.
05432251 INSPEC Abstract Number: A9315-4370-013, B9308-6130-067, C9308-5260S-004
Title: Efficient joint compensation of speech for the effects of additive noise and linear filtering
Author Liu, F.-H.; Acero, A.; Stern, R.M.
Author Affiliation: Dept. of Electr. & Comput. Eng., Carnegie Mellon Univ., Pittsburgh, PA, USA
Conference Title: ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech and Signal Processing (Cat. No.92CH3103-9) p. 257-60 vol.1
Publisher: IEEE , New York, NY, USA
Publication Date: 1992 Country of Publication: USA 5 vol. 3219 pp.
ISBN: 0 7803 0532 9
U.S. Copyright Clearance Center Code: 0 7803 0532 9/92/$3.00
Conference Sponsor: IEEE
Conference Date: 23-26 March 1992 Conference Location: San Francisco, CA, USA
Language: English
Subfile: A B C
Author Liu, F.-H.; Acero, A.; Stern, R.M.
Abstract: ...a fashion that is suitable for real-time environmental normalization for workstations of moderate size. The first algorithm is a modification of the SNR-dependent cepstral normalization (SDCN) and the fixed code-word dependent cepstral normalization (FCDCN) algorithms given by Acero and Stern (1990), except that unlike these algorithms it provides computationally-efficient environment normalization without prior knowledge of the...
...complexity, and amount of training data needed to adapt to new acoustical environments using these algorithms with several different types of headset-mounted and desktop microphones.
Identifiers: ...SNR-dependent cepstral normalization... ...fixed code-word dependent cepstral normalization...
Astronomical Objects:

24/3,K/2 (Item 2 from file: 2) Links
INSPEC
(c) 2008 Institution of Electrical Engineers. All rights reserved.
04864200 INSPEC Abstract Number: B91026317, C91029263
Title: Environmental robustness in automatic speech recognition
Author Acero, A.; Stern, R.M.
Author Affiliation: Dept. of Electr. & Comput. Eng., Carnegie Mellon Univ., Pittsburgh, PA, USA
Conference Title: ICASSP 90. 1990 International Conference on Acoustics, Speech and Signal Processing (Cat.

No.90CH2847-2) p. 849-52 vol.2
Publisher: IEEE , New York, NY, USA
Publication Date: 1990 Country of Publication: USA 5 vol. 2970 pp.
U.S. Copyright Clearance Center Code: CH2847-2/90/0000-0849$01.00
Conference Sponsor: IEEE
Conference Date: 3-6 April 1990 Conference Location: Albuquerque, NM, USA
Language: English
Subfile: B C
Author Acero, A.; Stern, R.M.
Abstract: ...system, robust to changes in the environment are reported. To deal with differences in noise level and spectral tilt between close-talking and desk-top microphones, two novel methods based on additive corrections in the cepstral domain are proposed. In the first algorithm, the additive correction depends on the instantaneous SNR of the signal. In the second technique, expectation-maximization techniques are used to best match the cepstral vectors of the input utterances to the ensemble of codebook entries representing a standard acoustical ambience. Use of the algorithms dramatically improves recognition accuracy when the system is tested on a microphone other than the one on which it was trained.
Descriptors: ...microphones;
Identifiers: close-talking microphones; ... ...desk-top microphones; cepstral domain

24/3,K/3 (Item 1 from file: 6) Links
Fulltext available through: Check for PDF Download Availability and Purchase
NTIS

2372625 NTIS Accession Number: ADA457727/XAB
Towards Environment-Independent Spoken Language Systems

Acero, A. ; Stern, R. M.
Carnegie-Mellon Univ., Pittsburgh, PA. School of Computer Science.
Corporate Source Codes: 005343049; 423887

1990 7p
Language: English
Journal Announcement: USGRDR0709
Product reproduced from digital image. Order this product from NTIS by: phone at 1-800-553-NTIS (U.S. customers); (703)605-6000 (other countries); fax at (703)605-6900; and email at orders@ntis.gov. NTIS is located at 5285 Port Royal Road, Springfield, VA, 22161, USA.
NTIS Prices: PC A02/MF A01
Acero, A. ; Stern, R. M.
...independent recognition system, robust to changes in the environment. To deal with differences in noise level and spectral tilt between close-talking and desk-top microphones, we describe two novel methods based on additive corrections in the cepstral domain. In the first algorithm, an additive correction is imposed that depends on the instantaneous SNR of the signal. In the second technique, EM techniques are used to best match the cepstral vectors of the input utterances to the ensemble of codebook entries representing a standard acoustical ambience. Use of these algorithms dramatically improves recognition accuracy when the system is tested on a microphone other than the one on which it was trained.
Descriptors: *Algorithms; *Speech recognition; *Microphones; Signal processing; Accuracy; Corrections; Noise; Additives; Words(Language)

24/3,K/4 (Item 2 from file: 6) Links
Fulltext available through: Check for PDF Download Availability and Purchase
NTIS

2371218 NTIS Accession Number: ADA458659/XAB
Efficient CEPSTRAL Normalization for Robust Speech Recognition
(Conference paper)
Liu, F. ; Stern, R. M. ; Huang, X. ; Acero, A.
Carnegie-Mellon Univ., Pittsburgh, PA. School of Computer Science.
Corporate Source Codes: 005343049; 423887

1993 7p
Language: English

Journal Announcement: USGRDR0708
Product reproduced from digital image. Order this product from NTIS by: phone at 1-800-553-NTIS (U.S. customers); (703)605-6000 (other countries); fax at (703)605-6900; and email at orders@ntis.gov. NTIS is located at 5285 Port Royal Road, Springfield, VA, 22161, USA.
NTIS Prices: PC A02/MF A01
Efficient CEPSTRAL Normalization for Robust Speech Recognition
Liu, F. ; Stern, R. M. ; Huang, X. ; Acero, A.
...environment- independent extension of the efficient SDCN and FCDCN algorithms developed previously. We compare the performance of these algorithms with the very simple RASTA and cepstral mean normalization procedures, describing the performance of these algorithms in the context of the 1992 DARPA CSR evaluation using secondary microphones, and in the DARPA stress-test evaluation.
Descriptors: *Language; *Speech recognition; *Telephone systems; *Acoustics; Algorithms; Environments; Models; Speech; Microphones; Office buildings; Quiet; Hearing; Gages; Passenger vehicles; Industrial plants; Secondary; Recognition; Floors; Arrays; Accuracy

24/3,K/5 (Item 3 from file: 6) Links
Fulltext available through: Check for PDF Download Availability and Purchase
NTIS

2371213 NTIS Accession Number: ADA458654/XAB
Acoustical Pre-Processing for Robust Speech Recognition
(Conference paper)
Stern, R. M. ; Acero, A.
Carnegie-Mellon Univ., Pittsburgh, PA. School of Computer Science.
Corporate Source Codes: 005343049; 423887

1989 9p
Language: English
Journal Announcement: USGRDR0708
Product reproduced from digital image. Order this product from NTIS by: phone at 1-800-553-NTIS (U.S. customers); (703)605-6000 (other countries); fax at (703)605-6900; and email at orders@ntis.gov. NTIS is located at 5285 Port Royal Road, Springfield, VA, 22161, USA.
NTIS Prices: PC A02/MF A01
Stern, R. M. ; Acero, A.
...to make SPHINX, the CMU continuous speech recognition system, environmentally robust. Our work has two major goals: to enable SPHINX to adapt to changes in microphone and acoustical environment, and to improve the performance of SPHINX when it is trained and tested using a desk-top microphone. This talk will describe some of our work in acoustical pre-processing techniques, specifically spectral normalization and spectral subtraction performed using an efficient pair of algorithms that operate primarily in the cepstral domain. The effects of these signal processing algorithms on the recognition accuracy of the Sphinx speech recognition system was compared using speech simultaneously recorded from two types of microphones: the standard close-talking Sennheiser HMD224 microphone and the desk-top Crown PZM6fs microphone. A naturally- elicited alphanumeric speech database was used. In initial results using the stereo alphanumeric database, we found that both the spectral subtraction and spectral normalization algorithms were able to provide very substantial improvements in recognition accuracy when the system was trained on the close-talking microphone and tested on the desk-top microphone, or vice versa. Improving the recognition accuracy of the system when trained and tested on the desk-top microphone remains a difficult problem requiring more sophisticated noise suppression techniques.
Descriptors: *Signal processing; *Speech recognition; *Acoustics; Data bases; Environments; Speech; Normalizing(Statistics); Alphanumeric data; Microphones; Noise reduction; Efficiency; Spectra ; Accuracy; Algorithms

24/3,K/6 (Item 1 from file: 8) Links
Ei Compendex(R)

08916343 E.I. No: EIP01436697119
Title: Efficient on-line acoustic environment estimation for FCDCN in a continuous speech recognition system

Author: Droppo, J.; Acero, A.; Deng, L.
Corporate Source: Microsoft Research One Microsoft Way, Redmond, WA 98052, United States
Conference Title: 2001 IEEE Interntional Conference on Acoustics, Speech, and Signal Processing
Conference Location: Salt Lake, UT, United States Conference Date: 20010507-20010511

E.I. Conference No.: 58541
Source: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings v 1 2001. p 209-212 (IEEE cat n 01CH37221)
Publication Year: 2001
CODEN: IPRODJ ISSN: 0736-7791
Language: English
Author: Droppo, J.; Acero, A.; Deng, L.
Abstract: There exists a number of cepstral de-noising algorithms which perform quite well when trained and tested under similar acoustic environments, but degrade quickly under mismatched conditions. We present two key...
Descriptors: *Speech recognition; Algorithms; Acoustic noise; Microphones; Signal to noise ratio; Acoustic signal processing
Identifiers: On-line acoustic environment estimation; Cepstral de-noising algorithms
Identifiers:

24/3,K/7 (Item 1 from file: 35) Links
Dissertation Abs Online

01161439 ORDER NO: AAD91-17502
ACOUSTICAL AND ENVIRONMENTAL ROBUSTNESS IN AUTOMATIC SPEECH RECOGNITION (ACOUSTICAL ROBUSTNESS, SPEECH RECOGNITION)

Author: ACERO, ALEJANDRO
Degree: PH.D.
Year: 1990
Corporate Source/Institution: CARNEGIE-MELLON UNIVERSITY (0041)
Source: Volume 5201B of Dissertations Abstracts International.
PAGE 413 . 156 PAGES
Author: ACERO, ALEJANDRO

...attempt to improve the recognition accuracy of speech recognition systems when they are trained and tested in different acoustical environments, and when a desk-top microphone (rather than a close-talking microphone) is used for speech input. Without such processing, mismatches between training and testing conditions produce an unacceptable degradation in recognition accuracy.

Two kinds of environmental variability are introduced by the use of desk-top microphones and different training and testing conditions: additive noise and spectral tilt introduced by linear filtering. An important attribute of the novel compensation algorithms described in... ...provide joint rather than independent compensation for these two types of degradation.

Acoustical compensation is applied in our algorithms as an additive correction in the cepstral domain. This allows a higher degree of integration within SPHINX, the Carnegie Mellon speech recognition system, that uses the cepstrum as its feature vector. Therefore... ...of vector-quantized cepstra that are produced by the feature extraction module in SPHINX.

In this dissertation we describe several algorithms including the SNR-Dependent Cepstral Normalization, (SDCN) and the Codeword-Dependent Cepstral Normalization (CDCN). With CDCN, the accuracy of SPHINX when trained on speech recorded with a close-talking microphone and tested on speech recorded with a desk-top microphone is essentially the same obtained when the system is trained and tested on speech from the desk-top microphone.

An algorithm for frequency normalization has also been proposed in which the parameter of the bilinear transformation that is used by the signal-processing stage...

[File 344] Chinese Patents Abs Jan 1985-2006/Jan

[File 347] JAPIO Dec 1976-2007/Dec(Updated 080328)

[File 350] Derwent WPIX 1963-2008/UD=200843

[File 371] French Patents 1961-2002/BOPI 200209


```
Set      Items   Description
S1      208764   S MIC OR MICS OR TRANSDUCER?? OR MICROPHONE??
S2         213   S MEL()FREQUENCY()CEPSTRAL()COEFFICIENT?? OR MFCC OR CEPTSTRUM
OR CEPSTRAL
S3        8895   S (WEIGHT??? OR SCORE?? OR SCORING OR NUMERIC? OR SUM?? OR
ADD?? OR SCALAR?? OR COEFFICIENT??)(5N)VECTOR??
S4      791641   S STATISTIC? OR STOCHASTIC? OR PROBABILITY OR LIKELY OR
LIKELIHOOD OR PROXIM? OR APPROXIMAT??? OR ESTIMAT???
S5       25808   S AU=(LIU, Z? OR LIU Z? OR SINCLAIR, M? OR SINCLAIR M? OR
ACERO, A? OR ACERO A? OR HUANG, X? OR HUANG X? OR DROPPO, J? OR DROPPO J? OR
DENG, L? OR DENG L? OR ZHANG, Z? OR ZHANG Z? OR ZHENG, Y? OR ZHENG Y?)
S6      167238   S (REDUCE?? OR MIN OR MINIMUM OR MINIMIS?? OR MINIMIZ??? OR
REDUCTION OR REDUCING OR FILTER??? OR ELIMINAT??? OR ATTENUAT???)(3N)(NOISE??
OR ARTIFACT??)
S7          15   S S1 AND S2
S8           3   S S7 AND S3
S9           2   S S8 AND S4
S10          0   S S7 AND S6
S11         11   S S7 AND IC=G10L?
S12          5   S S11 NOT (S8 OR AD=20031126:20080710/PR)
S13         89   S S3 AND S4 AND S6
S14          1   S S13 AND S2
S15          0   S S14 NOT (S8 OR S12 OR AD=20031126:20080710/PR)
S16          7   S S13 AND S1
S17          5   S S16 NOT (S8 OR S12 OR AD=20031126:20080710/PR)
S18          5   S S5 AND S13
S19          0   S S18 NOT (S8 OR S12 OR S17 OR AD=20031126:20080710/PR)
S20         16   S S13 AND IC=G10L?
S21          7   S S20 NOT (S8 OR S12 OR S17 OR AD=20031126:20080710/PR)
S22          1   S S8 NOT S9
```

9/3,K/1 (Item 1 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0012865984 & & *Drawing available*
WPI Acc no: 2002-724942/200279
XRPX Acc No: N2002-571621
Hidden Markov models modification method for speech recognition, involves adding mean cepstrum coefficient vector corresponding to speech to original HMM mean vectors
Patent Assignee: GONG Y (GONG-I); TEXAS INSTR INC (TEXI)
Inventor: GONG Y

Patent Family (7 patents, 28 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type

EP 1241662	A2	20020918	EP 2002100251	A	20020314	200279	B
US 20020173959	A1	20021121	US 2001275487	P	20010314	200279	E
			US 200251640	A	20020118		
JP 2002311989	A	20021025	JP 200267939	A	20020313	200303	E
US 7062433	B2	20060613	US 2001275487	P	20010314	200639	E
			US 200251640	A	20020118		
EP 1241662	B1	20060621	EP 2002100251	A	20020314	200643	E
DE 60212477	E	20060803	DE 60212477	A	20020314	200654	E
			EP 2002100251	A	20020314		
DE 60212477	T2	20070705	DE 60212477	A	20020314	200744	E
			EP 2002100251	A	20020314		

Priority Applications (no., kind, date): US 2001275487 P 20010314; US 200251640 A 20020118; EP 2002100251 A 20020314

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
EP 1241662	A2	EN	9	2		
Regional Designated States,Original	AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI TR					
US 20020173959	A1	EN			Related to Provisional	US 2001275487
JP 2002311989	A	JA	10			
US 7062433	B2	EN			Related to Provisional	US 2001275487
EP 1241662	B1	EN				
Regional Designated States,Original	DE FR GB					
DE 60212477	E	DE			Application	EP 2002100251
					Based on OPI patent	EP 1241662
DE 60212477	T2	DE			Application	EP 2002100251
					Based on OPI patent	EP 1241662

Hidden Markov models modification method for speech recognition, involves adding mean cepstrum coefficient vector corresponding to speech to original HMM mean vectors Alerting Abstract ...NOVELTY - The mean mel-scaled cepstrum coefficient (MFCC) vector corresponding to a speech, is calculated and added to the original hidden Markov models (HMM) mean vectors. An estimate of the background noise is determined. The mean vector of the noisy speech portion is determined, which is removed from the model mean vector corresponding... Original Publication Data by AuthorityArgentinaPublication No. Original Abstracts: A method of speech recognition with compensation is provided by modifying HMM models trained on clean speech with cepstral mean normalization. For each speech utterance the MFCC vector is calculated for the clean speech database. This mean MFCC is added to the original models. An estimate of the background noise is determined for a given speech utterance. The model mean vectors adapted to the noise are determined. The mean vector over... ... A method of speech recognition with compensation is provided by modifying HMM models trained on clean speech with cepstral mean normalization. For each spech utterance the MFCC vector is calculated for the clean database. This mean MFCC vector is added to the original models. An estimate of the background noise is determined for a given speech utterance. The model mean vectors adapted to the noise is determined. The mean vector of... ... A method of speech recognition with compensation is provided by modifying HMM models trained on clean speech with cepstral mean normalization. For all speech utterances the MFCC vector is calculated for the clean database. This mean MFCC vector is added to the original models. An estimate of the background noise is determined for a given speech utterance. The model mean vectors adapted to the noise are determined. The mean vector of... Claims: A method of modifying HMM models trained on clean speech with cepstral mean normalization to provide models that compensate for simultaneous channel/microphone distortion and background noise (additive distortion) comprising the steps of: for each speech utterance calculating the mean mel-scaled cepstrum coefficients (MFCC) vector [b.ABOVE.^]over the clean database; adding the mean MFCC vector [b.ABOVE.^]to the mean vectors m p,j,k of the original HMM models where p is the index of PDF, j is the state, and k the mixing component to get in m p,j,k ; for a given speech utterance calculating an estimate of the background noise vector X~ ; calculating the model mean vectors adapted to the noise X~ using [m.ABOVE.^] p,j,k = IDFT (DFT (m... ... the Inverse Discrete Fourier Transform is taken sum of the Discrete Fourier Transform of the mean vectors m p,j,k modified by the mean MFCC vector [b.ABOVE.^]added to the

Discrete Fourier Transform of the estimated noise X~ ; and calculating the mean vector [b.ABOVE.^]of the noisy data over the noisy speech space, and removing the mean vector [b.ABOVE... ... Verfahren zum Modifizieren originaler Hidden-Markov-Modelle (HMM), die an unverfalschter Sprache mit Cepstral-Mittelwertnomierung trainiert sind, um angepasste Modelle zu schaffen, die gleichzeitig eine Faltungsverzerrung und additives Rauschen kompensieren, das die folgenden Schritte umfasst: fur jede Sprachausserung: Berechnen (2) eines mittleren mel-skalierten Cepstrumkoeffizienten-Vektors (MFCC-Vektor) b uber die unverfalschte Datenbank;Addieren (3) des mittleren MFCC-Vektors b zu den mittleren Vektoren mp,j,k der originalen HMM-Modelle, wobei p der Index der Wahrscheinlichkeitsdichtefunktion (PDF) ist, j der Zustand ist... ... A method of modifying original Hidden Markov models (HMM) trained on clean speech with cepstral mean normalization to provide adapted models that compensate for simultaneous convolutive distortion and additive noise comprising the steps of: for each speech utterance: calculating (2) a mean mel-scaled cepstrum coefficients (MFCC) vector b over the clean database;adding (3) the mean MFCC vector b to the mean vectors mp,j,k of the original HMM models, where p is the index of the probability density function of (PDF), j is the state, and k the mixing component, to get mp,j,k;for the given speech utterance calculating (4) an estimate of the background noise vector X(tilde); calculating (5) model mean vectors adapted to the noise m(tilde)p,j,k = IDFT (DFT (m(tilde... ... combination (circled plus) of the Discrete Fourier Transform of the mean vectors m(circled plus)p,j,k and the Discrete Fourier Transform of the estimated noise X(tilde), the combination operator (circled plus) being defined by [MF IMGB0018] with [MF IMGB0019] and [MF IMGB0020] ; andcalculating (6) a mean vector ... et le bruit additif simultanes comprenant les etapes de: pour chaque enonce vocal: calcul (2) d'un vecteur de coefficients cepstraux en l'echelle mel moyen (MFCC) b sur la base de donnees pure;addition (3) du vecteur MFCC moyen b aux vecteurs moyens mp,j,k des modeles HMM originaux, ou p est l'indice de la fonction de densite de probabilite (PDF), j est l'etat et k la composante de melange, pour obtenir mep,j,k;pour l'enonce vocal donne, calcul (4) d'une estimation du vecteur de bruit de fond X(tilde); calcul (5) de vecteurs moyens de modele adaptes au bruit m^ p,J,k= IDFT(DFT... ... What is claimed is: 1. A method of modifying HMM models trained on clean speech with cepstral mean normalization to provide models that compensate for simultaneous channel/microphone distortion and background noise (additive distortion) comprising the steps of: for each speech utterance calculating the mean mel-scaled cepstrum coefficients (MFCC) vector [(circumflex).over.b] over the clean database; adding the mean MFCC vector [(circumflex).over.b] to the mean vectors mp,j,k of the original HMM models where p is the index of PDF, j is the state, and k the mixing component to get in mp,j,k; for a given speech utterance calculating an estimate of the background noise vector [(similar).over.X]; calculating the model mean vectors adapted to the noise [(similar).over.X] using mp,j,k... ... where the Inverse Discrete Fourier Transform is taken sum of the Discrete Fourier Transform of the mean vectors m p,j,k modified by the mean MFCC vector [(circumflex).over.b] added to the Discrete Fourier Transform of the estimated noise [(similar).over.X]; and calculating the mean vector [(circumflex).over.b] of the noisy data over the noisy speech space, and removing the mediv xhtml:class="heading">What is claimed is:16. A method of speech recognition with simultaneous compensation for both channel/microphone distortion and background noise comprising the steps of: modifying HMM models trained on clean speech with cepstral mean normalization;for all training speech utterances calculating the MFCC vector for a clean database;adding this mean MFCC vector to the original HMM models;estimating the background noise for a given speech utterance;determining the model mean vectors adapted to the noise;determining the mean vector of the noisy data over the noisy spee...

9/3,K/2 (Item 2 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0010372126 & & *Drawing available*
WPI Acc no: 2000-015099/200002
XRPX Acc No: N2000-011879
Speech recognition using machine to transform speech into written text
Patent Assignee: ENTROPIC CAMBRIDGE RES LAB LTD (ENTR-N); ENTROPIC LTD (ENTR-N); ODELL J (ODEL-I)
Inventor: ODELL J

Patent Family (7 patents, 83 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
EP 953971	A1	19991103	EP 1998303440	A	19980501	200002	B
WO 1999057716	A1	19991111	WO 1999GB1261	A	19990426	200002	E
AU 199936199	A	19991123	AU 199936199	A	19990426	200016	E
EP 1074018	A1	20010207	EP 1999918167	A	19990426	200109	E
			WO 1999GB1261	A	19990426		

US 6370505	B1	20020409	US 1999302370	A	19990430	200227	E
EP 1074018	B1	20040324	EP 1999918167	A	19990426	200422	E
			WO 1999GB1261	A	19990426		
DE 69915817	E	20040429	DE 69915817	A	19990426	200429	E
			EP 1999918167	A	19990426		
			WO 1999GB1261	A	19990426		

Priority Applications (no., kind, date): EP 1998303440 A 19980501

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
EP 953971	A1	EN	17	2		
Regional Designated States,Original	AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI					
WO 1999057716	A1	EN				
National Designated States,Original	AE AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA UG UZ VN YU ZA ZW					
Regional Designated States,Original	AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW NL OA PT SD SE SL SZ UG ZW					
AU 199936199	A	EN			Based on OPI patent	WO 1999057716
EP 1074018	A1	EN			PCT Application	WO 1999GB1261
					Based on OPI patent	WO 1999057716
Regional Designated States,Original	DE FI FR GB IT NL SE					
EP 1074018	B1	EN			PCT Application	WO 1999GB1261
					Based on OPI patent	WO 1999057716
Regional Designated States,Original	DE FI FR GB IT NL SE					
DE 69915817	E	DE			Application	EP 1999918167
					PCT Application	WO 1999GB1261
					Based on OPI patent	EP 1074018
					Based on OPI patent	WO 1990057716

Alerting Abstract ...NOVELTY - A microphone (10) is coupled via an acoustic processing unit (12) to a decoder (14), decoding input speech received via the processing unit, while a computer terminal (16) displays the decoded speech or uses the speech as input commands. A processor (20) compares the speech vector, defined by a set of cepstral parameters, to a number of possible output states stored in a memory (24). By determining which output state most closely matches the input speech vector, the processor effectively decodes the input speech. Each output state is represented by a mixture of probability density functions known as state mixture components, approximated by a weighted sum of a number of predetermined generic components. ...10 Microphone Original Publication Data by AuthorityArgentinaPublication No. ...Original Abstracts:output states are defined with each output state (j) being represented by a number of state mixture components (m). Each state mixture component is then approximated by a weighted sum of a number of predetermined generic components (x), allowing the likelihoods of each output states (j) corresponding to the input speech vector (or) to be determined... ... output states are defined with each output state (j) being represented by a number of state mixture components (m). Each state mixture component is then approximated by a weighted sum of a number of predetermined generic components (x), allowing the likelihoods of each output states (j) corresponding to the input speech vector (or) to be determined... ... output states are defined with each output state (j) being represented by a number of state mixture components (m). Each state mixture component is then approximated by a weighted sum of a number of predetermined generic components (x), allowing the likelihoods of each output states (j) corresponding to the input speech vector (or) to be determined... ... output states are defined with each output state (j) being represented by a number of state mixture components (m). Each state mixture component is then approximated by a weighted sum of a number of predetermined generic components (x), allowing the likelihoods of each output states (j) corresponding to the input speech vector (or) to be determined. ... ... sortie etant represente par un certain nombre de composantes (m) formant le melange d'etat. Chaque composante du melange d'etat est ensuite evaluee par approximation a l'aide de la somme ponderee d'un certain nombre de composantes (x) generiques

predeterminees, ce qui permet de determiner la vraisemblance de chaque etat (j) de sortie correspondant au vecteur (or) des signaux vocaux d'entree. ...Claims: von Zustands-Mischkomponenten (m) dargestellt wird, jede Zustands-Mischkomponente eine durch eine gewichtete Summe einer Anzahl von vorbestimmten generischen Komponenten (x) approximierte Wahrscheinlichkeitsverteilungsfunktion ist, die Approximation den Schritt des Bestimmens eines Wichtungs-Parameters (wjmx) fur jede generische Komponente (x) fur jede Zustands-Mischkomponente (m) enthalt unddas Verfahren zum Bestimmen der Wahrscheinlichkeiten des Ausgabezustands (j) die folgenden Schritte umfasst:1... ... input speech vector (or), wherein each output state (j) is represented by a number of state mixture components (m), each state mixture component being a probability distribution function approximated by a weighted sum of a number of predetermined generic components (x), the approximation including the step of determining a weighting parameter (wjmx) for each generic component (x) for each state mixture component (m),the method of determining the output state (j) likelihoods comprising the steps of:1) generating a correspondence probability signal representing a correspondence probability (Prx), wherein the correspondence probability (Prx) is the probability of each respective generic component (x) corresponding to the input speech vector (or);2) generating a threshold signal, representing a threshold value Tmix;3) selecting a number of output states (Nj);4) determining, for each state mixture component (m) of each selected output state (j), whether a weighted probability (gjmr) given by the scalar product of the weighting parameters (wjmx) and the respective correspondence probabilities (Prx), exceeds the threshold value Tmix; and,5) generating a set of output signals representing state likelihoods (bj) for each selected output state (j) by evaluating the likelihoods of all the state mixture components (m) of the respective selected output state (j) which have a weighted probability (gjmr) exceeding the threshold Tmix.Procede de traitement de la parole, le procede consistant a... ... nombre d'etats de sortie (j) possibles correspondant au vecteur de la parole d'entree (or), dans lequel chaque etat de sortie (j) est represente par un nombre de composantes d'etat mixtes (m), chaque composante d'etat mixte etant une fonction de repartition cumulative de probabilite approximee par une somme ponderee d'un nombre de composantes generiques (x) predeterminees, l'approximation comprenant l'etape consistant a determiner un parametre de ponderation (wjmx) pour chaque composante generique (x) pour chaque composante d'etat mixte (m),le procederespectives (Prx) depasse la valeur de seuil Tmix; et5) generer un ensemble de signaux de sortie representant les vraisemblances d'etat (bj) pour chaque etat de sortie (j) choisi par l'evaluation des vraisemblances de toutes... ... input speech vector (or), wherein each output state (j) is represented by a number of state mixture components (m), each state mixture component being a probability distribution function approximated by a weighted sum of a number of predetermined generic components (x), the approximation including the step of determining a weighting parameter (wjmx) for each generic component (x) for each state mixture component (m),
the method of determining the output state (j) likelihoods comprising the steps of:
1) generating a correspondence probability signal representing a correspondence probability (Prx), wherein the correspondence probability (Prx) is the probability of each respective generic component (x) corresponding to the input speech vector (or);
2) generating a threshold signal, representing a threshold value Tmix;
3) selecting a number of output states (Nj);
4) determining, for each state mixture component (m) of each selected output state (j), whether a weighted probability (gjmr) given by the scalar product of the weighting parameters (wjmx) and the respective correspondence probabilities (Prx), exceeds the threshold value Tmix; and,
5) generating... ... output state (j) by evaluating the likelihoods of all the state mixture components (m) of the respective selected output state (j) which have a weighted probability (gjmr) exceeding the threshold Tmix. A method of processing speech, the method comprising:receiving the speech and determining therefrom an input speech vector (or) representing a sample of the speech to be processed; and,determining the likelihoods of a number of possible output states (j) corresponding to the input speech vector (or), wherein each output state (j) is represented by a number of state mixture components (m), each state mixture component being a probability distribution function approximated by a weighted sum of a number of predetermined generic components (x), the approximation including the step of determining a weighting parameter (wjmx) for each generic component (x) for each state mixture component (m),the method of determining the output state (j) likelihoods comprising the steps of:1) generating a correspondence probability signal representing a correspondence probability (Prx), wherein the correspondence probability (Prx) is the probability of each respective generic component (x) corresponding to the input speech vector (or);2) generating a threshold signal, representing a threshold value Tmix;3) selecting a number of output states (Nj);4) determining, for each state mixture component (m) of each selected output state (j), whether a weighted probability (gjmr) given by the scalar product of the weighting parameters (wjmx) and the respective correspondence probabilities (Prx), exceeds the threshold value Tmix; and,5) generating a set of output signals representing state likelihoods (bj)for each selected output state (j) by evaluating the likelihoods of all the state mixture components (m) of the respective selected output state (j) which have a weighted probability (gjmr) exceeding the threshold Tmix.

12/3,K/1 (Item 1 from file: 347) Links
Fulltext available through: Order File History
JAPIO

07443478 **Image available**

SPEECH RECOGNITION METHOD WITH CORRECTED CHANNEL DISTORTION AND CORRECTED BACKGROUND NOISE

Pub. No.: 2002-311989 [JP 2002311989 A]
Published: October 25, 2002 (20021025)
Inventor: GONG YIFAN
Applicant: TEXAS INSTRUMENTS INC
Application No.: 2002-067939 [JP 200267939]
Filed: March 13, 2002 (20020313)
Priority: 01 275487 [US 2001275487], US (United States of America), March 14, 2001 (20010314)
International Class: G10L-015/14; G10L-015/02; G10L-015/06; G10L-015/20; G10L-021/02

ABSTRACT

PROBLEM TO BE SOLVED: To provide a model for correcting the distortion and background noise of a channel/microphone at the same time.

SOLUTION: The speech recognition method with correction is provided by correcting a hidden Markov model trained in a clean voice by cepstral mean normalization. An MFCC vector is calculated to a clean voice database with respect to each voice utterance. Its mean MFCC is added to an original model. An estimate of background noise is determined with respect to a given voice utterance. A model mean vector applied... Di01

12/3,K/2 (Item 1 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0012906288 & & *Drawing available*
WPI Acc no: 2002-229223/200229
XRPX Acc No: N2002-176198
Speech recognition system for mobile communication devices is capable of rapidly processing greater varieties of words and operable in many different devices
Patent Assignee: VERBALTEK CO LTD (VERB-N); VERBALTEK INC (VERB-N)
Inventor: CHANG J; CHEN J; CHEN J Y; KIM Y; PAN J

Patent Family (7 patents, 30 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
EP 1139332	A2	20011004	EP 2000309816	A	20001106	200229	B
CN 1315809	A	20011003	CN 2000109843	A	20000707	200229	E
EP 1139332	A9	20020320				200229	E
KR 2001096490	A	20011107	KR 200060110	A	20001012	200229	E
US 6304844	B1	20011016	US 2000538657	A	20000330	200229	E
JP 2002108387	A	20020410	JP 200153125	A	20010123	200240	E
TW 504663	A	20021001	TW 2001106813	A	20010322	200337	E

Priority Applications (no., kind, date): US 2000538657 A 20000330

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
EP 1139332	A2	EN	19	10		
Regional Designated States,Original	AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI TR					
EP 1139332	A9	EN				
Regional Designated States,Original	AL AT BE CH CY DE DK ES FI FR GB GR IE IT LI LT LU LV MC MK NL PT RO SE SI TR					
JP 2002108387	A	JA	64			
TW 504663	A	ZH				

Class Codes International Patent Classification IPC Class Level Scope Position Status Version Date G10L-015/00 Main G10L-0015/02... ...G10L-0015/10... ...G10L-0015/18... ...G10L-0015/22 G10L-0015/00... Original Publication Data by AuthorityArgentinaPublication No. ...Original Abstracts:to-machine communication for mobile phones, PDAs, and other communication devices. This invention is an apparatus and method for a speech recognition system comprising a microphone, front-end signal processor for generating parametric representations of speech input signals, a pronunciation database, a letter similarity comparator for comparing the parametric representation of the input signals with... ... to-machine communication for mobile phones, PDAs, and other communication devices. This invention is an apparatus and method for a speech recognition system comprising a microphone, front-end signal processor for generating parametric representations of speech input signals, a pronunciation database, a letter similarity comparator for comparing the parametric representation of the input signals with the... Claims: A speech recognition system comprising: microphone means for receiving acoustic waves and converting the acoustic waves into electronic signals; front-end signal processing means, coupled to said microphone means, for processing the electronic signals to generate parametric representations of the electronic signals; pronunciation database storage means for storing a plurality of parametric representations of letter pronunciations; letter similarity comparator... ... A speech recognition system comprising: microphone means for receiving acoustic waves and converting the acoustic waves into electronic signals; front-end signal processing means, coupled to said microphone means, for processing the electronic signals to generate parametric representations of the electronic signals, including preemphasizer means for spectrally flattening the electronic signals generated by said microphone means; frame-blocking means, coupled to said preemphasizer means, for blocking the electronic signals into frames of N samples with adjacent frames separated by M samples; windowing means, coupled to said frame-blocking means, for windowing each frame; autocorrelation means, coupled to said windowing means, for autocorrelating the frames; cepstral coefficient generating means, coupled to said autocorrelation means, for converting each frame into cepstral coefficients; and tapered windowing means, coupled to said cepstral coefficient generating means, for weighting the cepstral coefficients, thereby generating parametric representations of the sound waves; pronunciation database storage means for storing a plurality of parametric representations of letter pronunciations; letter similarity comparator means, coupled to said front-end signal processing means and to said pronunciation database storage means, for comparing the parametric representation of the electronic signals...

12/3,K/3 (Item 2 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0012486680 & & *Drawing available*
WPI Acc no: 2002-433836/200246
XRPX Acc No: N2002-341382
Speech recognition system has subtractor that subtracts utterance log spectral mean from log spectral vector of incoming speech signal, and recognizer that receives output of subtractor
Patent Assignee: TEXAS INSTR INC (TEXI)
Inventor: GONG Y; RAMALINGAM C S

Patent Family (1 patents, 1 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
US 6381571	B1	20020430	US 199883926	P	19980501	200246	B
			US 1999293357	A	19990416		

Priority Applications (no., kind, date): US 199883926 P 19980501; US 1999293357 A 19990416

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
US 6381571	B1	EN	7	7	Related to Provisional	US 199883926

Class Codes International Patent Classification IPC Class Level Scope Position Status Version Date G10L-0015/06... ...G10L-0015/20 G10L-0015/00... Original Publication Data by AuthorityArgentinaPublication No. Original Abstracts:Utterance-based mean removal in log-domain, or in any linear transformation of log-domain, e.g., cepstral domain, is known to improve substantially a recognizerprimes robustness to transducer difference, channel distortion,

and speaker variation. Applicants teach a sequential determination of utterance log-spectral mean by a generalized maximum a posteriori estimation. The solution...

12/3,K/4 (Item 3 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0006703448
WPI Acc no: 1994-084290/199411
XRPX Acc No: N1994-065945
Normalising adaptive speech recognition system - subtracting mean cepstral vector component value from each cepstral vector component
Patent Assignee: DAIMLER-BENZ AG (DAIM); DAIMLERCHRYSLER AG (DAIM)
Inventor: CLASS F; KALTENMEIER A; REGEL-BRIETZMANN P

Patent Family (5 patents, 2 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
DE 4229577	A1	19940310	DE 4229577	A	19920904	199411	B
EP 586996	A2	19940316	EP 1993113800	A	19930828	199411	E
EP 586996	A3	19940413	EP 1993113800	A	19930828	199522	E
EP 586996	B1	19991020	EP 1993113800	A	19930828	199948	E
DE 59309837	G	19991125	DE 59309837	A	19930828	200002	E
			EP 1993113800	A	19930828		

Priority Applications (no., kind, date): DE 4229577 A 19920904

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
DE 4229577	A1	DE	5	1		
EP 586996	A2	DE	5			
Regional Designated States,Original	DE FR GB IT					
EP 586996	A3	EN				
EP 586996	B1	DE				
Regional Designated States,Original	DE FR GB IT					
DE 59309837	G	DE			Application	EP 1993113800
					Based on OPI patent	EP 586996

...subtracting mean cepstral vector component value from each cepstral vector component Alerting Abstract ...The speech recognitions system is normalised with a long term spectrum for adaption to the microphone and speech characteristics by subtracting the mean value of the cepstral vector component from the cepstral vector component. The obtained normalisation eliminates the time variant and quasi-time variant noise signals... ...Pref. the mean value of the cepstral vector component is adjusted via an adaption factor to obtain slow tracking of the mean value under relatively stable conditions... ...ADVANTAGE - Allows speech recognition system to be modified for each different speaker or microphone. Class Codes International Patent Classification IPC Class Level Scope Position Status Version Date G10L-003/00... ...G10L-005/06 Main Original Publication Data by AuthorityArgentinaPublication No. ...Original Abstracts:invention relates to a method for normalising a speech recognition system to a long-time spectrum in order to eliminate interfering influences such as different microphone, room and speaker characteristics even before the classifier is reached. The cepstral features which are used in known speech recognition systems are normalised by subtracting the mean value of the cepstral vector components from the respective cepstral vector. The method is advantageously expanded by adaptive normalisation by sliding mean-value estimation with a sliding window with respect to time and an adaptive time constant. This makes... ...Claims:1. Method for standardising a speech-adaptive speech recognition system to a long-term spectrum, characterised in that
- the cepstral features of a speech-recognition system are used,
- the mean values of the cepstral vector components are subtracted from the cepstral vector components and the cepstral features are thereby standardised for any speech-recognition system, and

- the mean values of the cepstral vector components are smoothly adapted by way of a drag window to changing conditions of speech and environment.

12/3,K/5 (Item 4 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0006567982 & & *Drawing available*
WPI Acc no: 1993-379392/199348
XRPX Acc No: N1993-293017
Voice recognition method using neuronal network - involves recognising pronounce words by comparison with words in reference vocabulary using sub-vocabulary for acoustic word reference
Patent Assignee: SOLLAC SA (SOLL-N)
Inventor: ANGLADE Y; FOHR D; HENRYON M; STOUFFLET F

Patent Family (1 patents, 1 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
FR 2689292	A1	19931001	FR 19923743	A	19920327	199348	B

Priority Applications (no., kind, date): FR 19923743 A 19920327

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes
FR 2689292	A1	FR	23	7	

Alerting Abstract ...The method involves using a voice recognition circuit (11) using, for example, Markov chains, a detection circuit (12), a discriminatory selection circuit (13), a cepstral analyzer (14) and one (or several) neuronal networks. A word spoken into a microphone is sent as numeric signal and transformed into acoustic wave patterns. These patterns are then compared with data characteristic of predetermined reference words... ...the word does not fit one of the identifiers then it is retransmitted through the circuit (12) to output (S1) of the recognition system. A cepstral analysis on the waveforms by analyser (14) produces temporal filtration and is transformed by Fourier and inverse transforms. Synoptic coefficients with under-vocabulary association are... Class Codes International Patent Classification IPC Class Level Scope Position Status Version Date G10L-0015/16... G10L-0015/00.

17/3,K/1 (Item 1 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0017224487 & & *Drawing available*
WPI Acc no: 2008-A44917/200803
XRPX Acc No: N2008-033994
Tap weight selection method for adaptive multi-tap frequency domain digital filter for beam forming transducer array in radar, involves applying robustness-control transformation to reduce target canceling components of vector
Patent Assignee: MASSACHUSETTS INST TECHNOLOGY (MASI)
Inventor: DESLODGE J G

Patent Family (1 patents, 1 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
US 7254199	B1	20070807	US 1998100291	P	19980914	200803	B
			US 1999396175	A	19990914		

Priority Applications (no., kind, date): US 1998100291 P 19980914; US 1999396175 A 19990914

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
US 7254199	B1	EN	54	17	Related to Provisional	US 1998100291

Tap weight selection method for adaptive multi-tap frequency domain digital filter for beam forming transducer array in radar, involves applying robustness-control transformation to reduce target canceling components of vector Original Titles:Location-estimating, null steering (LENS) algorithm for adaptive array processing Alerting Abstract ...NOVELTY - Each tap weight is parameterized such that each tap weight is characterized by vector of parameters. Each parameter is solved by minimizing the expected power of the array output signal. A robustness-control transformation is applied to the vector to identify and reduce target canceling components of vector that arise from incomplete target location knowledge while preserving non-target canceling components. The weight vector indicative of the filter tap weights is formed as a function of the vector. ...USE - For adaptive multi-tap frequency domain digital filter for beam forming transducer array used in radar, sonar, satellite communications, and background noise reducing hearing aids... Title Terms .../Index Terms/Additional Words: TRANSDUCER; Class Codes Original Publication Data by AuthorityArgentinaPublication No. Original Abstracts:An adaptive multiple-tap frequency domain digital filter processes an input signal vector X from an plurality of spatially separated transducers that detect energy from a plurality of sources including a target energy source and at least one non-target energy source. The filter receives and processes the input signal vector X to attenuate noise from non-target sources and provides an output signal vector Y. Tap weights WN for the filter are selected by first parameterizing each of the tap weights WN, such that each of the tap weights WN is characterized by a vector of parameters betaopt, and the solving for each parameter of the vector betaopt by minimizing the expected power of the array output signal... ... opt wherein the robustness-control transformation identifies and reduces target canceling components of the vector betaopt while preserving non-target canceling components. Finally, the weight vector indicative of the filter tap weights is formed as a function of the vector betaopt. Notably, the present invention separates the robustness constraining process from the beamforming power minimization, in... ...Claims:tap weights WN for an adaptive multi-tap frequency domain digital filter that processes an input signal vector X from a plurality of spatially separated transducers that detect energy from a plurality of sources including a target energy source and at least one non-target energy source, wherein the filter receives and processes the input signal vector X to attenuate noise from non-target sources and provides an output signal vector Y, the method comprising the steps of:parameterizing each of the tap weights WN such that each of the tap weights WN is characterized by a vector of parameters betaopt;solving for each parameter of the vector betaopt by minimizing the expected power of the array output signal Y;applying... ... reduces target canceling components of the vector betaopt that arise from incomplete target location knowledge while preserving non-target canceling components; andforming the weight vector indicative of the filter tap weights as a function of the vector betarob.

17/3,K/2 (Item 2 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0014642009 & & *Drawing available*
WPI Acc no: 2004-824028/200482
XRPX Acc No: N2004-650810
Ultrasonic transmit and receive pulse optimizing method for ultrasonic imaging, involves applying weighting parameters and delays of minimum of energy function to signals for exciting transducers to produce comprehensional pulse
Patent Assignee: CURLETTO S (CURL-I); ESAOTE SPA (ESAO-N); TRUCCO A (TRUC-I)
Inventor: CURLETTO S; TRUCCO A

Patent Family (3 patents, 34 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
EP 1480199	A1	20041124	EP 2004101916	A	20040504	200482	B
US 20050033167	A1	20050210	US 2004852572	A	20040524	200512	E
US 7297117	B2	20071120	US 2004852572	A	20040524	200778	E

Priority Applications (no., kind, date): IT 2003SV23 A 20030522

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes

EP 1480199	A1	EN	22	15		
Regional Designated States,Original	AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IT LI LT LU LV MC MK NL PL PT RO SE SI SK TR					

Ultrasonic transmit and receive pulse optimizing method for ultrasonic imaging, involves applying weighting parameters and delays of minimum of energy function to signals for exciting transducers to produce comprehensional pulse Alerting Abstract ...profiles, and between ideal and actual beam patterns, is defined. Parameters and delays relating to minimum of the function are applied to signals for exciting transducers to produce a comprehensional ultrasonic pulse. ... ADVANTAGE - The method obtains better pressure profiles as well as an effective reduction of artifacts and a better angular resolution. The method thus improves ultrasound imaging performances while reducing computational load required for determining optimization weights... ... DESCRIPTION OF DRAWINGS - The drawing shows a schematic view of an array of transducers, propagation geometry of an ultrasonic pulse that is focused on a scan line in a pulse propagation direction within a body under examination.Title Terms .../Index Terms/Additional Words: TRANSDUCER; Class Codes Original Publication Data by AuthorityArgentinaPublication No. ...Original Abstracts:transmit and receive ultrasound pulses, particularly for ultrasonic imaging, wherein transmit pulses are generated from ultrasonic pulse contributions of each of a plurality of electroacoustic transducers, which are grouped in an array and are individually energized by electric excitation signals, an excitation signal being applied to each individual transducer of the array with a predetermined delay with respect to the application of the excitation signal to the other transducers, and a weight being applied to the excitation signal of each transducer for increasing/decreasing the amplitude of the excitation signal and, as a result, of the acoustic signal generated by the transducer. The invention envisages to optimize at least the amplitude weights of the individual transducers' contributions, by defining an energy function and by minimizing it. By using stochastic or evolutionary algorithms, the minimization of the energy function provides a weight vector for the contributions of the individual transducers, which leads to transmit or receive pulses having a predetermined minimum distance from desired pulses. As described regarding transmission, optimization may be also applied to... ... A method for optimizing transmit and receive ultrasound imaging pulses generates transmit pulses from an array of transducers which are energized by excitation signals that are applied to each individual transducer of the array. Each of the excitation signals are individually weighted to optimize the transducers' contribution to a predetermined energy function. Such optimization may also be performed on the received pulses... ... A method for optimizing transmit and receive ultrasound imaging pulses generates transmit pulses from an array of transducers which are energized by excitation signals that are applied to each individual transducer of the array. Each of the excitation signals are individually weighted to optimize the transducers' contribution to a predetermined energy function. Such optimization may also be performed on the received pulses. ...Claims:ultrasonic transmit and receive pulses, particularly for ultrasound imaging, wherein transmit pulses are generated from ultrasonic pulse contributions of each of a plurality of electroacoustic transducers, which are grouped in an array and are individually triggered by electric excitation signals, the excitation signal being applied to each individual transducer of the array with a predetermined delay with respect to the application of the excitation signal to the other transducers, and a weight being applied to the excitation signal of each transducer for increasing/decreasing the amplitude of the excitation signal and, as a result, the acoustic signal generated by the transducer, characterized in that it includes the following steps: at least for the transmit pulse, defining the optimal desired mechanical pressure profile, relative to the penetration depth of the ultrasonic pulse within the body or object being examined, as a function at least of amplitude weighting parameters for transducers' contributions to the comprehensive pulse, and of the delays of excitation for transmission of individual pulse contributions of transducers, aimed at focusing the ultrasonic pulse on a scan line or band and at a certain penetration depth within the body or object under examination... ... relative to the propagation time or penetration depth within the body or object under examination, as a function at least of amplitude weighting parameters for transducers' contributions to the comprehensive pulse, and of delays of excitation delays for transmission of individual pulse contributions of transducers aimed at focusing the ultrasonic pulse on a scan line or band and at a certain penetration depth within the body or object under examination... ... delays which correspond to the minimum of the energy function, applying said weighting parameters and said delays at least to the signals for exciting the transducers to generate the comprehensive ultrasonic pulse... ... method for optimizing ultrasonic pulses, particularly for ultrasound imaging, wherein transmit pulses are generated from ultrasonic pulse contributions of each of a plurality of electroacoustic transducers, said transducers being grouped in an array and being individually triggered by electric excitation signals, said excitation signal being applied to each individual transducer of said array having a predetermined delay with respect to the application of the excitation signal that is applied to the other transducers of said plurality of transducers, and wherein a weight is applied to the excitation signal for each transducer for adjusting the amplitude of said excitation signal, characterized in the following steps: defining an optimal desired mechanical pressure profile for said transmit pulses relative... ... the penetration depth of said transmit pulses within the body or object being examined as a function of at least amplitude weighting parameters for said transducers' contributions to said transmit pulses, and of the delays of excitation for transmission of individual pulse contributions of transducers, aimed at focusing comprehensive pulses on a scan line or band and at a certain penetration depth within the body or

object under examination; defining... ... to the propagation time or penetration depth within the body or object under examination as a function of at least amplitude weighting parameters for said transducers' contributions to said transmit pulses, and of delays of excitation delays for transmission of individual pulse contributions of transducers aimed at focusing comprehensive pulses on a scan line or band and at a certain penetration depth within the body or object under examination; defining... ... delays which correspond to the minimum of the energy function and applying said weighting parameters and said delays to said excitation signals for exciting said transducers to generate said comprehensive pulses... ... or more ultrasonic pulses in conjunction with ultrasonic imaging, wherein transmit pulses are generated from ultrasonic pulse contributions of each of a plurality of electroacoustic transducers, said transducers being grouped in an array and being individually triggered by electric excitation signals, said excitation signal being applied to each individual transducer of said array having a predetermined delay with respect to the application of the excitation signal that is applied to the other transducers of said plurality of transducers, and wherein a weight is applied to the excitation signal for each transducer for adjusting the amplitude of said excitation signal, characterized in the following steps: defining an optimal desired mechanical pressure profile for said transmit pulses relativeof at least amplitude weighting parameters for said transducers' contributions to said transmit pulses, and of the delays of excitation for transmission of individual pulse contributions of transducers, aimed at focusing comprehensive pulses on a scan line or band and at a certain penetration depth within the body or object under examination;defining... ... to the propagation time or penetration depth within the body or object under examination as a function of at least amplitude weighting parameters for said transducers' contributions to said transmit pulses, and of delays of excitation delays for transmission of individual pulse contributions of transducers aimed at focusing comprehensive pulses on a scan line or band and at a certain penetration depth within the body or object under examination;defining... ... delays which correspond to the minimum of the energy function and applying said weighting parameters and said delays to said excitation signals for exciting said transducers to generate said comprehensive pulses.

17/3,K/3 (Item 3 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0013733586 & & *Drawing available*
WPI Acc no: 2003-831696/200377
Related WPI Acc No: 2006-008205; 2006-724192
XRPX Acc No: N2003-664651
Noise reduction in speech recognition system, involves adding input feature vector obtained by converting noise signal frame, and correction vector selected based on feature vector, to obtain clean feature vector
Patent Assignee: ACERO A (ACER-I); DENG L (DENG-I); DROPPO J G (DROP-I); MICROSOFT CORP (MICT)
Inventor: ACERO A; DENG L; DROPPO J G

Patent Family (2 patents, 1 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
US 20030191638	A1	20031009	US 2002117142	A	20020405	200377	B
US 7117148	B2	20061003	US 2002117142	A	20020405	200665	E

Priority Applications (no., kind, date): US 2002117142 A 20020405

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes
US 20030191638	A1	EN	20	11	

Noise reduction in speech recognition system, involves adding input feature vector obtained by converting noise signal frame, and correction vector selected based on feature vector, to obtain... Original Titles:Method of noise reduction using correction vectors based on dynamic aspects of speech and noise normalization... ...Method of noise reduction using correction vectors based on dynamic aspects of speech and noise normalization Alerting Abstract ...noise signal is converted into input feature vector, based on which a correction vector which incorporates dynamic aspects of pattern signals, is selected. The correction vector is added to the input feature vector to obtain clean feature vector. DESCRIPTION - An INDEPENDENT CLAIM is also included for computer readable medium storing noise reduction program... ...ADVANTAGE - Improves the performance of noise reduction system... ...DESCRIPTION OF DRAWINGS - The figure shows the block diagram of a computing system for implementing noise reduction. ... ...163 microphone Original Publication Data by AuthorityArgentinaPublication No. Original Abstracts:A method and

apparatus are provided for reducing noise in a signal. Under one aspect of the invention, a correction vector is selected based on a noisy feature vector that represents a noisy signal. The selected correction vector incorporates dynamic aspects of pattern signals. The selected correction vector is then added to the noisy feature vector to produce a cleaned feature vector. In other aspects of the invention, a noise value is produced from an estimate of the noise in a noisy signal. The noise value is subtracted from a value representing a portion of the noisy signal to produce a... ... A method and apparatus are provided for reducing noise in a signal. Under one aspect of the invention, a correction vector is selected based on a noisy feature vector that represents a noisy signal. The selected correction vector incorporates dynamic aspects of pattern signals. The selected correction vector is then added to the noisy feature vector to produce a cleaned feature vector. In other aspects of the invention, a noise value is produced from an estimate of the noise in a noisy signal. The noise value is subtracted from a value representing a portion of the noisy signal to produce a... Claims:What is claimed is:1. A method for reducing noise in a noisy input signal, the method comprising:converting a frame of the noisy input signal into an input feature vector;selecting a mixture component... ... What is claimed is:1. A method for reducing noise in a noisy input signal, the method comprising: converting a frame of the noisy input signal into an input feature vector;selecting a mixture component... ... on the input feature vector;identifying a correction vector that incorporates dynamic aspects of a pattern signal based on the selected mixture component, the correction vector having at least one delta coefficient; andadding the correction vector to the input feature vector to form a clean feature vector.

17/3,K/4 (Item 4 from file: 350) Links

Fulltext available through: Order File History

Derwent WPIX

0013085360 & & *Drawing available*

WPI Acc no: 2003-165973/200316

XRPX Acc No: N2003-131070

Robust pattern recognition method for handheld devices, involves combining specific set of feature vectors satisfying predetermined equation to obtain optimized set of feature vectors to represent specific pattern

Patent Assignee: INT BUSINESS MACHINES CORP (IBMC)

Inventor: GAO Y; PICHENY M A; RAMABHADRAN B

Patent Family (2 patents, 1 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
US 20020152069	A1	20021017	US 2000238841	P	20001006	200316	B
			US 2001968051	A	20011001		
US 7054810	B2	20060530	US 2001968051	A	20011001	200636	E

Priority Applications (no., kind, date): US 2000238841 P 20001006; US 2001968051 A 20011001

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
US 20020152069	A1	EN	19	7	Related to Provisional	US 2000238841

Alerting Abstract ...vectors is preselected. The sets are combined to obtain optimized set of vectors satisfying an equation including an exponential function (Sj), with respect to conditional probability P(X1,X2-XN/Sj) for feature vectors and weights (W1,W2-Wn) assigned to feature vectors. ... ADVANTAGE - A speech recognition includes increased recognition accuracy, robustness to car/background noise, reduced computational costs, ability to merge with other streams including non-audio streams such as video, and reduced memory usage... Original Publication Data by AuthorityArgentinaPublication No. ...Original Abstracts:in a manner to obtain an optimized set of feature vectors which best represents the pattern. The combination is performed via one of a weighted likelihood combination scheme and a rank-based state-selection scheme; preferably, it is done in accordance with an equation set forth herein. In one aspect, a weighted likelihood combination can be employed, while in another aspect, rank-based state selection can be employed. An apparatus suitable for performing the method is described, and implementation in a computer... ... sets of feature vectors are combined in a manner to obtain an optimized set of feature vectors which best represents the pattern. The combination is performed via one of a weighted likelihood combination scheme and a rank-based state-selection scheme; preferably, it is done in accordance with an equation set forth herein. In one aspect, a weighted likelihood combination can be employed, while in another aspect, rank-based state selection can be employed. An apparatus suitable for performing the method is described, and implementation in a computer program product is also

contemplated... ...Claims:function log( ),sj is a label for a class j,N is greater than or equal to 2,p(x1, x2,. .. xN|sj) is conditional probability of feature vectors x1, x2,. .. xN given that they are generated by said class j,K is a normalization constant,w1, w2,. .. wN are weights... ... x2,. .. xN from a set of observation vectors which are indicative of a pattern of an analog input signal converted to electronic form by a transducer which it is desired to recognize, at least one of said sets of feature vectors being different than at least one other of said sets of feature vectors and being preselected for purposes of containing at least some complimentary information with regard to said at least one other of said sets of feature vectors; and(b) combining said N sets of feature vectors in a manner to obtain an optimized set of feature vectors which best represents said pattern... ... function log( ),sj is a label for a class j,N is greater than or equal to 2,p(x1, x2,. .. xN|sj) is conditional probability of feature vectors x1, x2,. .. xN given that they are generated by said class j,K is a normalization constant,w1, w2,. .. wN are weights assigned to x1, x2,. .. xN respectively according to confidence levels therein; andq is a real number corresponding to a desired combination function.

17/3,K/5 (Item 5 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0012988804 & & *Drawing available*
WPI Acc no: 2003-066519/200306
XRPX Acc No: N2003-051579
Digital signal filter for audio processing application, estimates echo portion of microphone signal based on real and imaginary frequency coefficients corresponding to speaker and microphone signals
Patent Assignee: MICROSOFT CORP (MICT)
Inventor: MALVAR H S

Patent Family (1 patents, 1 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
US 6473409	B1	20021029	US 1999259390	A	19990226	200306	B

Priority Applications (no., kind, date): US 1999259390 A 19990226

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes
US 6473409	B1	EN	24	14	

Digital signal filter for audio processing application, estimates echo portion of microphone signal based on real and imaginary frequency coefficients corresponding to speaker and microphone signals Original Titles:Adaptive filtering system and method for adaptively canceling echoes and reducing noise in digital signals Alerting Abstract ...NOVELTY - A modulated complex lapped transform processor spectrally decomposes the speaker and microphone signals at predefined frequencies, into corresponding real and imaginary frequency coefficients. An adaptive filter receives the frequency coefficients and computes an output signal, that is an estimate of an echo portion of the microphone signal. ... Adaptive echo cancellation device; and Noise reduction device. ... ... ADVANTAGE - Allows a perfect signal reconstruction, as echo cancellation is easily performed, by estimating the microphone signals echo portion.Title Terms .../Index Terms/Additional Words: ESTIMATE; ... ...MICROPHONE; Class Codes Original Publication Data by AuthorityArgentinaPublication No. ...Original Abstracts:produce resulting real and imaginary vectors, respectively. The real and imaginary transform processors compute spatial transforms on the real and imaginary vectors to produce real and imaginary transform coefficient of the MCLT, respectively. The MCLT is a biorthogonal spectral transformation system, in the sense that the original time domain signal can be reconstructed exactly..

21/3,K/1 (Item 1 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0013918528 & & *Drawing available*
WPI Acc no: 2004-098290/200410
XRPX Acc No: N2004-078377

Speech coding system for communication system, computes peak to average ratio of linear prediction spectrum, based on spectrum sampling frequency of frame, and broadens LP filter coefficients, based on ratio of LP spectrum
Patent Assignee: BHASKAR U (BHAS-I); SWAMINATHAN K (SWAM-I)
Inventor: BHASKAR U; SWAMINATHAN K

Patent Family (1 patents, 1 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
US 20040002856	A1	20040101	US 2002362706	P	20020308	200410	B
			US 2003382202	A	20030305		

Priority Applications (no., kind, date): US 2002362706 P 20020308; US 2003382202 A 20030305

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
US 20040002856	A1	EN	59	12	Related to Provisional	US 2002362706

Class Codes International Patent Classification IPC Class Level Scope Position Status Version Date G10L-0019/08... G10L-0019/00... Original Publication Data by AuthorityArgentinaPublication No. ...Original Abstracts:2 Kbps. At 4 Kbps, the codec uses a 20 ms frame size and a 20 ms lookahead for purposes of voice activity detection (VAD), noise reduction, linear prediction (LP) analysis, and open loop pitch analysis. The LP parameters are encoded using backward predictive hybrid scalar-vector quantizers in the line spectral frequency (LSF) domain after adaptive bandwidth broadening to minimize excessive peakiness in the LP spectrum. Prototype Waveforms (PW) are extracted every subframe or 2... ... correlations and voicing measure. The phase component is generated based on a first order vector autoregressive model in which each PW vector is generated by summing the previous PW vector weighted by the decoded PW correlation coefficient with a weighted combination of a fixed and random phase components. The use of the PW correlations in this manner results in a sequence of PWs that exhibit... ... by tilt correction and energy normalization. At 2.4 Kbps, the same frame size of 20 ms and a lookahead of 20 ms for VAD, noise reduction, LP analysis, and pitch estimation are utilized. However, the LP parameters are encoded using a 3-stage 21 bit VQ with backward prediction. Furthermore, for encoding the PW parameters an additional 20 ms of... ...Claims:prediction (LP) front end, adapted to process an input signal which provides LP parameters that are computed during a predetermined interval;an open loop pitch estimator, adapted to perform pitch frequency estimation on said input signal for substantially all of said predetermined intervals;an adaptive bandwidth broadening module, adapted to perform the following operations:derive a spectrum sampling frequency for said predetermined interval as the pitch frequency or its integer...

21/3,K/2 (Item 2 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0010637021 & & *Drawing available*
WPI Acc no: 2001-244032/200125
Related WPI Acc No: 2001-137489; 2001-137490; 2001-218097; 2003-265468; 2003-742688
XRPX Acc No: N2001-173750
Frequency domain interpolative coding system for low bit-rate coding of speech signals, uses code book of preset content and operates on decimated prototype waveform gain vector obtained from low pass filter
Patent Assignee: HUGHES ELECTRONICS CORP (HUGA)
Inventor: BHASKAR B R U; NANDKUMAR S; SWAMINATHAN K; UDAYA BHASKAR B R; ZAKARIA G

Patent Family (4 patents, 82 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
WO 2000060579	A1	20001012	WO 2000US8790	A	20000404	200125	B
AU 200041902	A	20001023	AU 200041902	A	20000404	200125	E
EP 1088304	A1	20010404	EP 2000921609	A	20000404	200126	E
			WO 2000US8790	A	20000404		
US 6418408	B1	20020709	US 1999127780	P	19990405	200370	E
			US 2000542792	A	20000404		

Priority Applications (no., kind, date): US 1999127780 P 19990405; US 2000542792 A 20000404

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
WO 2000060579	A1	EN	82	7		
National Designated States,Original	AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GE GH GM HR HU ID IL IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA UG UZ VN YU ZW					
Regional Designated States,Original	AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW NL OA PT SD SE SZ TZ UG ZW					
AU 200041902	A	EN			Based on OPI patent	WO 2000060579
EP 1088304	A1	EN			PCT Application	WO 2000US8790
					Based on OPI patent	WO 2000060579
Regional Designated States,Original	AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE					
US 6418408	B1	EN	32		Related to Provisional	US 1999127780

Alerting Abstract ...to input signal, provides linear prediction (LP) parameters which are quantized and encoded over predetermined intervals and computes LP residual signal. An open loop pitch estimator (24) in response to the LP residual signal yields a pitch contour within the predetermined interval... ...24 Open loop pitch estimator Class Codes International Patent Classification IPC Class Level Scope Position Status Version Date G10L-0011/02... ...G10L-0011/02... ...G10L-0011/04... ...G10L-0019/00... ...G10L-0019/02... ...G10L-0019/04... ...G10L-0019/08... ...G10L-0019/14 G10L-0011/00... ...G10L-0011/00... ...G10L-0019/00 Original Publication Data by AuthorityArgentinaPublication No. ...Original Abstracts:by a piecewise-constant construction. The REW phase vector is regenerated at the decoder based on the received REW gain vector and the voicing measure, which determines a weighted mixture of SEW component and a random noise that is passed through a high pass filter to generate the REW component. The high pass filter poles are adjusted based on the voicing measure to control the REW component characteristics. At the... ...Claims:input signal providing LP parameters which are quantized and encoded over predetermined intervals and used to compute a LP residual signal;an open loop pitch estimator responsive to said LP residual signal, a pitch quantizer, and a pitch interpolator yielding a pitch contour within the predetermined interval;a signal processor responsive to said LP residual signal and...

21/3,K/3 (Item 3 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0010535011 & & *Drawing available*
WPI Acc no: 2001-137490/200114
Related WPI Acc No: 2001-137489; 2001-218097; 2001-244032; 2003-265468; 2003-742688
XRPX Acc No: N2001-100166
Frequency domain interpolative coding system for magnitude modeling of prototype waveforms, has fixed dimension vector quantizers to suitably quantize variable dimension slowly evolving waveform magnitude vector
Patent Assignee: HUGHES ELECTRONICS CORP (HUGA)
Inventor: BHASKAR B R U; NANDKUMAR S; SWAMINATHAN K; UDAYA BHASKAR B R; ZAKARIA G

Patent Family (4 patents, 82 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
WO 2000060578	A1	20001012	WO 2000US8995	A	20000405	200114	B
AU 200041978	A	20001023	AU 200041978	A	20000405	200114	E
EP 1133767	A1	20010919	EP 2000921699	A	20000405	200155	E
			WO 2000US8995	A	20000405		
US 6493664	B1	20021210	US 1999127780	P	19990405	200326	E
			US 2000542793	A	20000404		

Priority Applications (no., kind, date): US 1999127780 P 19990405; US 2000542793 A 20000404

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
WO 2000060578	A1	EN	82	7		
National Designated States,Original	AL AM AT AU AZ BA BB BG BR BY CA CH CN CU CZ DE DK EE ES FI GB GE GH GM HR HU ID IL IS JP KE KG KP KR KZ LC LK LR LS LT LU LV MD MG MK MN MW MX NO NZ PL PT RO RU SD SE SG SI SK SL TJ TM TR TT UA UG UZ VN YU ZW					
Regional Designated States,Original	AT BE CH CY DE DK EA ES FI FR GB GH GM GR IE IT KE LS LU MC MW NL OA PT SD SE SL SZ TZ UG ZW					
AU 200041978	A	EN			Based on OPI patent	WO 2000060578
EP 1133767	A1	EN			PCT Application	WO 2000US8995
					Based on OPI patent	WO 2000060578
Regional Designated States,Original	AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE					
US 6493664	B1	EN	31		Related to Provisional	US 1999127780

Alerting Abstract ... response to an input signal provides LP parameters which are quantized and encoded over predetermined intervals, to compute LP residual signal. An open loop pitch estimator (24) in response to the LP residual signal, pitch quantizer and pitch interpolator yields a pitch contour within the preset interval. The signal processor in... ... 24 Open loop pitch estimator Class Codes International Patent Classification IPC Class Level Scope Position Status Version Date G10L-0011/02... ...G10L-0011/02... ...G10L-0011/04... ...G10L-0019/00... ...G10L-0019/02... ...G10L-0019/04... ...G10L-0019/14 G10L-0011/00... ...G10L-0011/00... ...G10L-0019/00 Original Publication Data by AuthorityArgentinaPublication No. ...Original Abstracts:the REW magnitude is explicitly coded. The REW phase vector is regenerated at the decoder based on the received REW gain vector and the voicing measure, which determines a weighted mixture of SEW component and a random noise that is passed through a high pass filter to generate the REW component. The high pass filter poles are adjusted based on the voicing measure to control the REW component characteristics. At the... ... and only the REW magnitude is explicitly coded. The REW phase vector is regenerated at the decoder based on the received REW gain vector and the voicing measure, which determines a weighted mixture of SEW component and a random noise that is passed through a high pass filter to generate the REW component. The high pass filter poles are adjusted based on the voicing measure to control the REW component characteristics. At the ouput filter, the magnitude of the REW component is... ...Claims:input signal providing LP parameters which are quantized and encoded over predetermined intervals and used to compute a LP residual signal;an open loop pitch estimator responsive to said LP residual signal;a pitch quantizer;a pitch interpolator;said open loop pitch estimator, said pitch quantizer, and said pitch interpolator yielding a pitch contour within the predetermined interval;a signal processor responsive to said LP residual signal and the pitch contour for extracting a prototype waveform (PW) for a number of equal sub-intervals within...

21/3,K/4 (Item 4 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0009824663 & & *Drawing available*
WPI Acc no: 2000-115703/200010
Related WPI Acc No: 2000-363415
XRPX Acc No: N2000-087567
Noise reduction for speech recognition system
Patent Assignee: ADVANCED MICRO DEVICES INC (ADMI)
Inventor: ASGHAR S M; CONG L

Patent Family (1 patents, 1 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
US 6009391	A	19991228	US 1997883980	A	19970627	200010	B
			US 1997907145	A	19970806		

Priority Applications (no., kind, date): US 1997883980 A 19970627; US 1997907145 A 19970806

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
US 6009391	A	EN	19	3	C-I-P of application	US 1997883980

Noise reduction for speech recognition system Alerting Abstract USE - Noise reduction in speech recognition... ...ADVANTAGE - By using a fuzzy matrix and energy coefficients the noise elements can be reduced Class Codes International Patent Classification IPC Class Level Scope Position Status Version Date G10L-0015/02... ...G10L-0015/10... ...G10L-0015/20 G10L-0015/00... Original Publication Data by AuthorityArgentinaPublication No. ...Original Abstracts:extraction followed by a fuzzy matrix quantizer (FMQ). Frames of the speech input signal are represented in a matrix by a vectorf of line spectral pair frequencies and energy coefficients and are fuzzy matrix quantized to respective vector f^ entries of a matrix codeword in a codebook of the FMQ. The energy coefficients include the original energy and the first and second derivatives... ... and a predicted speech input signal at the ith line spectral pair frequency of the speech input signal, the first G LSP frequencies are most likely to be frequency shifted by noise, and the last P+3 coefficients represent the three energy coefficients. This robust distance measure can be used to enhance speech recognition performance... ...Claims:determining P order line spectral pair frequencies for the speech input signal; representing the energy coefficients and line spectral pair frequencies as components of a vector: determining respective differences between the energy coefficients of the speech input signal and corresponding energy coefficients of a plurality of reference codewords; determining respective differences between the respective P line spectral frequencies of the speech input signal and corresponding P line...

21/3,K/5 (Item 5 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0007184032 & & *Drawing available*
WPI Acc no: 1995-226601/199530
XRPX Acc No: N1995-177544
Swirl artifact removal system for CELP based speech coder - removes low frequency components of encoder input when non-periodic signal e.g. noise is detected
Patent Assignee: HUGHES AIRCRAFT CO (HUGA); HUGHES ELECTRONICS (HUGA)
Inventor: GANESAN K; GUPTA P; LEE H

Patent Family (7 patents, 16 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
EP 660301	A1	19950628	EP 1994850222	A	19941212	199530	B
CA 2136891	A	19950621	CA 2136891	A	19941129	199538	E
FI 199405915	A	19950621	FI 19945915	A	19941215	199538	E
EP 660301	B1	19960605	EP 1994850222	A	19941212	199627	E
DE 69400229	E	19960711	DE 69400229	A	19941212	199633	E
			EP 1994850222	A	19941212		
US 5633982	A	19970527	US 1993169789	A	19931220	199727	E
			US 1996734210	A	19961021		
CN 1113586	A	19951220	CN 1994112982	A	19941219	199739	E

Priority Applications (no., kind, date): US 1993169789 A 19931220; US 1996734210 A 19961021

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
EP 660301	A1	EN	9	4		
Regional Designated States,Original	AT BE CH DE DK ES FR GB GR IT LI NL SE					
CA 2136891	A	EN				
EP 660301	B1	EN	11	4		

Regional Designated States,Original	AT BE CH DE DK ES FR GB GR IT LI NL SE					
DE 69400229	E	DE			Application	EP 1994850222
					Based on OPI patent	EP 660301
US 5633982	A	EN	8	4	Continuation of application	US 1993169789

Title Terms .../Index Terms/Additional Words: VECTOR-SUM Class Codes International Patent Classification IPC Class Level Scope Position Status Version Date G10L-009/14 Main "Version 7" G10L-005/00 G10L-0019/00... ...G10L-0019/12... ...G10L-0021/02 G10L-0019/00... ...G10L-0021/00 Original Publication Data by AuthorityArgentinaPublication No. ...Claims:signal containing periodic and non-periodic signals; a high pass filter also connected to receive the input signal and operable to remove low frequency components likely to cause the production of swirl artifacts from the input signal, the switch being controllable to selectively supply the input signal or an output of the high pass filter to the CELP based encoder; and a detector connected to receive the input signal and information from the CELP based encoder and generate... ... and to connect the output of the high pass filter to the CELP based encoder when non-periodic signals are detected; wherein low frequency components likely to cause the production of swirl artifacts are alternately filtered from the CELP based encoder input signal to thereby prevent the production of swirl artifacts.

21/3,K/6 (Item 6 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0006920056 & & *Drawing available*
WPI Acc no: 1994-316479/199439
XRPX Acc No: N1994-248588
Noise suppressor using code conversion table - obtains code by vector-quantising cepstrum coeffts. extracted from voice signal having added noise, and converts into code for voice with noise suppressed
Patent Assignee: SONY CORP (SONY)
Inventor: AKABANE M; KATO Y; WATARI M

Patent Family (1 patents, 1 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
US 5353408	A	19941004	US 1992998724	A	19921230	199439	B

Priority Applications (no., kind, date): JP 199218478 A 19920107

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes
US 5353408	A	EN	9	3	

Alerting Abstract ...generates two corresp. codes which are based respectively on vector-quantised feature parameters of the two voice signals. A code converter associates in terms of probability, the two codes, and converts the second code to the first code... ...ADVANTAGE - Uses adaptive inverse filtering to cancel noise component. Class Codes International Patent Classification IPC Class Level Scope Position Status Version Date G10L-0011/02... ...G10L-0015/04... ...G10L-0021/02... ...G10L-0021/04 G10L-0011/00... ...G10L-0015/00... ...G10L-0021/00 Original Publication Data by AuthorityArgentinaPublication No. ...Original Abstracts:a code of a voice with noise added thereto and a code of a voice without noise are associated with each other in terms of probability, is referred to in a code converter. Using the code converter, a code is obtained in a vector quantizer by vector-quantizing cepstrum coefficients extracted from the voice with noise added thereto, and is converted into a code of a voice obtained by suppressing the noise in the voice with noise added thereto. Linear predictive coefficients... Claims: A noise suppressor apparatus for reducing noise accompanying a spoken voice comprising: input means for providing an analog electrical signal corresponding to the spoken voice, said electrical signal including a component corresponding to said... ... according to recursive relationships, said predictive filter calculating a residual signal based on said first digital signal and said first LPCs; code generating means for vector-quantizing said cepstrum coefficients according to first and second code tables stored in memory to provide first codes associated with said cepstrum coefficients, said

first code table being formulated from a... ... code converting means for providing second codes based on said first codes according to a code conversion table stored in memory; decoder means for inverse vector-quantizing cepstrum coefficients vector quantized with said code generating means; a linear predictive calculator for calculating second LPCs according to cepstrum coefficients inverse vector-quantized by said decoder means; synthesis filter means for providing a second digital signal corresponding to said spoken voice, said synthesis filter means calculating said second digital...

21/3,K/7 (Item 7 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0005594326 & & *Drawing available*
WPI Acc no: 1991-202094/199128
XRPX Acc No: N1991-154572
Sound synthesising in recursive filter - vector quantising each set of coeffts. for all blocks to obtain clustered representative coefft. set and determn. transition probability
Patent Assignee: RAYTHEON CO (RAYT)
Inventor: DEAETT M A; DEAETT M A N; ROSENSTRAC P A; ROSENSTRACH P A

Patent Family (6 patents, 6 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
EP 436397	A	19910710	EP 1990314403	A	19901228	199128	B
CA 2031965	A	19910703				199137	E
US 5119425	A	19920602	US 1990459420	A	19900102	199225	E
			US 1991780836	A	19911023		
EP 436397	A3	19920902	EP 1990314403	A	19901228	199338	E
EP 436397	B1	19961030	EP 1990314403	A	19901228	199648	E
DE 69029030	E	19961205	DE 69029030	A	19901228	199703	E
			EP 1990314403	A	19901228		

Priority Applications (no., kind, date): US 1990459420 A 19900102; US 1991780836 A 19911023

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
EP 436397	A	EN				
Regional Designated States,Original	DE FR GB IT					
CA 2031965	A	EN				
US 5119425	A	EN	15		Continuation of application	US 1990459420
EP 436397	A3	EN				
EP 436397	B1	EN	23	11		
Regional Designated States,Original	DE FR GB IT					
DE 69029030	E	DE			Application	EP 1990314403
					Based on OPI patent	EP 436397

...vector quantising each set of coeffts. for all blocks to obtain clustered representative coefft. set and determn. transition probability Alerting Abstract ...Synthesis of the pseudo-random signal is provided by randomly selecting according to a cumulative transition probability, the cluster representative of a next successive block given the selected cluster representative of the previous block, the coefficient of each block time being applied to a noise-excited recursive filter to generate the pseudo-random synthesised signal... ...with a random number from a random number generator (63) triggered by the timing pulse. The comparator (82) thereby provides a cluster number and cumulative probability for block (K+1). The cluster coefficient memory (70) loads an LPC recursive filter (84) with LPC coefficients... Equivalent Alerting Abstract ...The pseudo-random or transient synthesized signal is provided by analysis of a number of related signals by vector quantization of linear predictive coding coefficients of time blocks of the signals and provides cumulative probability matrices for the transition from one cluster representative for one

block to a cluster representative of the next successive block of each of the signals... ...Synthesis of the pseudo-random signal is provided by randomly selecting according to a cumulative transition probability, the cluster representative of a next successive block given the selected cluster representative of the previous block. The coefft of each block time is applied to a noise-excited recursive filter to generate the pseudo-random synthesized signal. Synthesis includes probabilistic models using Markov transitions, to produce transient sounds such as sonar, hatch closings, and hull... Technology Focus Title Terms .../Index Terms/Additional Words: PROBABILITY Class Codes International Patent Classification IPC Class Level Scope Position Status Version Date "Version 7" ...G10L-0019/08 ...G10L-0019/00 Original Publication Data by AuthorityArgentinaPublication No. Original Abstracts: A pseudo-random synthesized signal is provided by analysis of a plurality of related signals by vector quantization of linear predictive coding coefficients (cluster representatives) of time blocks of the signals and providing cumulative probability matrices for the transition from one cluster representative for one block to a cluster representative of the next successive block of each of the signals. Synthesis of the pseudo-random signal is provided by randomly selecting according to a cumulative transition probability, the cluster representative of a next successive block given the selected cluster representative of the previous block, the coefficient of each block time being applied to a noise-excited recursive filter to generate the pseudo-random synthesized signal. A timing pulse generator (81) produces a pulse every T milliseconds, where T is the block time.A... ... with a random number from a random number generator (63) triggered by the timing pulse. The comparator (82) thereby provides a cluster number and cumulative probability for block (K+1). The cluster coefficient memory (70) loads an LPC recursive filter (84) with LPC coefficients... ... A pseudo-random or transient synthesized signal is provided by analysis of a plurality of related signals by vector quantization of linear predictive coding coefficients (cluster representatives) of time blocks of the signals and providing cumulative probability matrices for the transition from one cluster representative for one block to a cluster representative of the next successive block of each of the signals. Synthesis of the pseudo-random signal is provided by randomly selecting according to a cumulative transition probability, the cluster representative of a next successive block given the selected cluster representative of the previous block, the coefficient of each block time being applied to a noise-excited recursive filter to generate the pseudo-random synthesized signal. Synthesis includes probabalistic models using Markov transitions, to produce transient sounds such as sonar, hatch closings, and hull... ...Claims:block of samples to determine the linear prediction coding coefficient sets of a recursive filter for each of said blocks for each of said signals; vector quantizing each set of said coefficients for all of said blocks to obtain a clustered representative coefficient set for each of said blocks, each cluster representative coefficient set representing an entire cluster of vector quantized linear prediction coding coefficient sets; determining the probability of a transition from one cluster representative for one block to a next cluster representative for the next successive block for all the blocks of each of said signals;providing a cumulative probability value for each transition for corresponding blocks of each signal; storing said cumulative transition probability values and said cluster representative coefficients; successively generating a probability value for each successive block of the signal to be synthesized; providing a cluster representative of a block to read out of a memory of a corresponding set of cumulative probability values for a next successive block; comparing said successively generated probability value for said next successive block with said cumulative probability values stored for said next successive block to determine a selected cluster representative for said next successive block; providing a set of cluster representative coefficients reading from a memory containing said cluster representative coefficients corresponding to said selected cluster representative;providing said selected cluster coefficients to a noise-excited recursive filter for a time corresponding to said block; and repeating the above process for successive blocks to provide said synthesized signal... ... Synthesis of the pseudo-random signal is provided by randomly selecting according to a cumulative transition probability, the cluster representative of a next successive block given the selected cluster representative of the previous block, the coefficient of each block time being applied to a noise-excited recursive filter to generate the pseudo-random synthesised signal... ... with a random number from a random number generator (63) triggered by the timing pulse. The comparator (82) thereby provides a cluster number and cumulative probability for block (K+1). The cluster coefficient memory (70) loads an LPC recursive filter (84) with LPC coefficients... ... selecting of and the supplying to the filter (84) of each set of LPC coefficients (ao, ..., an) comprises the steps of: generating (81,63) a probability value; comparing (62) the generated probability value with a selected set of cumulative probability values (Pij) derived from the transition probabilities between the stored sets, for the desired signal and determining the cumulative probability value (Pij) indicated by the comparison in accordance with a predetermined criterion; selecting a respective one of the stored sets (70) corresponding to the indicated cumulative probability value (Pij), each of the cumulative probability values (Pij) being associated with a respective one of the stores sets (70); applying the selected stored set (ao, ..., an) to the recursive filter (84) for the duration of a block time (T); and selecting, in dependence upon the said indicated cumulative probability value (Pij), a next set of cumulative probability values (Pij) for comparison with a next generated (63) probability value.

S22/3,K/1 (Item 1 from file: 350) Links
Fulltext available through: Order File History
Derwent WPIX

0017089443 & & *Drawing available*

WPI Acc no: 2007-804400/200775
XRPX Acc No: N2007-639129
Perturbed phonetic string generating method for use in e.g. speech recognition, involves applying perturbation to feature vector set, and phonetically decoding perturbed feature vector set for producing perturbed phonetic string
Patent Assignee: MOTOROLA INC (MOTI)
Inventor: MA C C

Patent Family (3 patents, 117 & countries)

Patent Number	Kind	Date	Application Number	Kind	Date	Update	Type
US 20070239444	A1	20071011	US 2006277793	A	20060329	200775	B
WO 2007117814	A2	20071018	WO 2007US63752	A	20070312	200775	E
WO 2007117814	A3	20080522	WO 2007US63752	A	20070312	200835	E

Priority Applications (no., kind, date): US 2006277793 A 20060329

Patent Details

Patent Number	Kind	Lan	Pgs	Draw	Filing Notes	
US 20070239444	A1	EN	12	4		
WO 2007117814	A2	EN				
National Designated States,Original	AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE EG ES FI GB GD GE GH GM GT HN HR HU ID IL IN IS JP KE KG KM KN KP KR KZ LA LC LK LR LS LT LU LY MA MD MG MK MN MW MX MY MZ NA NG NI NO NZ OM PG PH PL PT RO RS RU SC SD SE SG SK SL SM SV SY TJ TM TN TR TT TZ UA UG US UZ VC VN ZA ZM ZW					
Regional Designated States,Original	AT BE BG BW CH CY CZ DE DK EA EE ES FI FR GB GH GM GR HU IE IS IT KE LS LT LU LV MC MT MW MZ NA NL OA PL PT RO SD SE SI SK SL SZ TR TZ UG ZM ZW					
WO 2007117814	A3	EN				
National Designated States,Original	AE AG AL AM AT AU AZ BA BB BG BR BW BY BZ CA CH CN CO CR CU CZ DE DK DM DZ EC EE EG ES FI GB GD GE GH GM GT HN HR HU ID IL IN IS JP KE KG KM KN KP KR KZ LA LC LK LR LS LT LU LY MA MD MG MK MN MW MX MY MZ NA NG NI NO NZ OM PG PH PL PT RO RS RU SC SD SE SG SK SL SM SV SY TJ TM TN TR TT TZ UA UG US UZ VC VN ZA ZM ZW					
Regional Designated States,Original	AT BE BG BW CH CY CZ DE DK EA EE ES FI FR GB GH GM GR HU IE IS IT KE LS LT LU LV MC MT MW MZ NA NL OA PL PT RO SD SE SI SK SL SZ TR TZ UG ZM ZW					

Alerting Abstract ...NOVELTY - The method involves generating a feature vector set e.g. Mel Frequency Cepstral Coefficient, from a spoken utterance e.g. voice tags. A perturbation is applied to the feature vector set for producing a perturbed feature vector set. A randomly distributed noise is added to the perturbed feature vector set and multiplied by a variance. The perturbed feature vector set is phonetically decoded for producing a perturbed phonetic string, where the phonetic string is... ... 102 Microphone

[File 9] Business & Industry(R) Jul/1994-2008/Jul 07

**File 9: UD names have been reset to reflect currency. All data is present.*

[File 15] ABI/Inform(R) 1971-2008/Jul 10

[File 16] Gale Group PROMT(R) 1990-2008/Jul 01

**File 16: Because of updating irregularities, the banner and the update (UD=) may vary.*

[File 20] Dialog Global Reporter 1997-2008/Jul 10

[File 47] Gale Group Magazine DB(TM) 1959-2008/Jun 27

[File 75] TGG Management Contents(R) 86-2008/Jun W4

[File 80] TGG Aerospace/Def.Mkts(R) 1982-2008/Jul 02

[File 88] Gale Group Business A.R.T.S. 1976-2008/Jun 18

[File 112] UBM Industry News 1998-2004/Jan 27

[File 148] Gale Group Trade & Industry DB 1976-2008/Jun 20

**File 148: The CURRENT feature is not working in File 148. See HELP NEWS148.*

[File 160] Gale Group PROMT(R) 1972-1989

[File 275] Gale Group Computer DB(TM) 1983-2008/Jul 02

[File 264] DIALOG Defense Newsletters 1989-2008/Jul 09

[File 484] Periodical Abs Plustext 1986-2008/Jun W4

[File 570] Gale Group MARS(R) 1984-2008/Jul 02

[File 620] EIU:Viewswire 2008/Jul 08

[File 621] Gale Group New Prod.Annou.(R) 1985-2008/Jun 19

[File 623] Business Week 1985-2008/Jul 10

[File 624] McGraw-Hill Publications 1985-2008/Jul 10

**File 624: Homeland Security & Defense and 9 Platt energy journals added Please see HELP NEWS624 for more*

[File 634] San Jose Mercury Jun 1985-2008/Jun 29

[File 635] Business Dateline(R) 1985-2008/Jul 09

[File 636] Gale Group Newsletter DB(TM) 1987-2008/Jul 02

[File 647] CMP Computer Fulltext 1988-2008/Jun W3

[File 696] DIALOG Telecom. Newsletters 1995-2008/Jul 09

[File 674] Computer News Fulltext 1989-2006/Sep W1

**File 674: File 674 is closed (no longer updates).*

[File 810] Business Wire 1986-1999/Feb 28

[File 813] PR Newswire 1987-1999/Apr 30

[File 587] Jane`s Defense&Aerospace 2008/Jul W1


```
Set      Items   Description
S1      324904   S MIC OR MICS OR TRANSDUCER?? OR MICROPHONE??
S2         776   S MEL()FREQUENCY()CEPSTRAL()COEFFICIENT?? OR MFCC OR CEPTSTRUM
OR CEPSTRAL
S3       12190   S (WEIGHT??? OR SCORE?? OR SCORING OR NUMERIC? OR SUM?? OR
ADD?? OR SCALAR?? OR COEFFICIENT??)(5N)VECTOR??
S4    23540064   S STATISTIC? OR STOCHASTIC? OR PROBABILITY OR LIKELY OR
LIKELIHOOD OR PROXIM? OR APPROXIMAT??? OR ESTIMAT???
S5        4658   S AU=(LIU, Z? OR LIU Z? OR SINCLAIR, M? OR SINCLAIR M? OR
ACERO, A? OR ACERO A? OR HUANG, X? OR HUANG X? OR DROPPO, J? OR DROPPO J? OR
DENG, L? OR DENG L? OR ZHANG, Z? OR ZHANG Z? OR ZHENG, Y? OR ZHENG Y?)
S6      123260   S (REDUCE?? OR MIN OR MINIMUM OR MINIMIS?? OR MINIMIZ??? OR
REDUCTION OR REDUCING OR FILTER??? OR ELIMINAT??? OR ATTENUAT???)(3N)(NOISE??
OR ARTIFACT??)
S7           2   S S1(S)S2
S8          37   S S1 AND S2
S9           8   S S8 AND S3
S10          8   S S9 AND S4
S11          6   S S10 NOT PY>2003
S12          6   S S11 NOT S7
S13          5   S S12 NOT PY>2003
S14          0   S S8 AND S5
S15          7   S S3(S)S4(S)S6
S16          6   RD  (unique items)
S17          0   S S16(S)(S1 OR S2)
S18          1   S S16 AND (S1 OR S2)
S19          1   S S18 NOT (S7 OR S13 OR PY>2003)
```

7/3,K/1 (Item 1 from file: 88) Links
Gale Group Business A.R.T.S.

03689737 Supplier Number: 17470941
PET studies of auditory and phonological processing: effects of stimulus characteristics and task demands.
Fiez, J.A.; Raichle, M.E.; Miezin, F.M.; Petersen, S.E.; Tallal, P.; Katz, W.F.
Journal of Cognitive Neuroscience , v7 , n3 , p357(19)
Summer , 1995
ISSN: 0898-929X
Language: English Record Type: Fulltext

Word Count: 11996 Line Count: 00991

```
...was read 4 times. From this total of 24 productions, the most clearly
audible tokens were
selected.
     Recording was done with a Shure SM-57 microphone placed 5 inches from the
lips, and a
Tascam 34-B tape recorder. These tape-recorded productions were then filtered
at 8 kHz, and
digitally sampled at 16,000 samples/sec. Spectral parameters were obtained
using an autoregression,
linear predictive coding (LPC) model, with cepstral-based fundamental frequency
extraction.
These parameters were used to create highly natural synthetic speech tokens.
The initial durations
of the vocoded stimuli ranged from 261...
```

7/3,K/2 (Item 1 from file: 484) Links
Periodical Abs Plustext

06960255
African elephant vocal communication II: rumble variation reflects the individual identity and emotional state of callers

Soltis, Joseph; Leong, Kirsten; Savage, Anne
Animal Behaviour (IANB) , v70 , p 589-599
Sep 2005
ISSN: 0003-3472 Journal Code: IANB
Document Type: Feature
Language: English Record Type: Abstract

Abstract:


```
...rumbles from six adult female African elephants housed at Disney's Animal
Kingdom (Lake Buena
Vista, Florida, U.S.A.). Subjects wore collars outfitted with microphones and
radiotransmitters that allowed recording of vocalizations from identified
individuals. Rumble
vocalizations were digitized and both source and filter features were measured
for each call...

...identity of the caller. Second, rumbles varied as a function of negative
emotional arousal. When
associating with dominant animals, subordinate females produced rumbles with
lower cepstral
coefficients, suggesting low tonality and unstable pitch in the voice, compared
to rumbles produced
outside of the presence of dominant animals. Rumbles as a whole...
```


13/3,K/1 (Item 1 from file: 9) Links
Business & Industry(R)

01985734 Supplier Number: 25473792 (USE FORMAT 7 OR 9 FOR FULLTEXT)
Frontier Design targets high-volume, command-and-control apps -- Speech-recognition core to open new markets
(Frontier Design is introducing a new low-cost speech-recognition core with better than 97% accuracy; device recognizes speech in variety of languages)

Electronic Engineering Times , p 38
October 25, 1999
Document Type: Journal ISSN: 0192-1541 (United States)
Language: English Record Type: Fulltext
Word Count: 964 (USE FORMAT 7 OR 9 FOR FULLTEXT)

TEXT:

...in a number of "giveaway" applications. "Talk to your postcard, and it will respond," said Mark Bloemandael, Frontier's applications manager in the Netherlands. More likely, it will find a place in speech-controlled car radios, an application that demands noise cancellation, high-accuracy speech recognition and low cost all in...

...a voice-controlled currency translation device for Columns Ltd. of Singapore. The design includes a complete speech recognition and synthesis (SRS) system on a chip, microphone and speaker. It includes 30 kbytes of RAM for the storage of speech recognition templates, and an additional 20 kbytes of ROM for the storage...

...other on-chip logic) or as a cell-based SoC including codec and amplifiers. Alternatively, Frontier can market a complete OEM module that includes speaker, microphone, IFR, RF and other functionality.

Since the speech recognition algorithm requires only 5 to 10 Mips, any existing pager, mobile telephone or other system with...

...of spare processing power can include the SRS core with no extra overhead.

The SRS implements several advanced recognition algorithms: the Mel Frequency Cepstrum Coefficient (MFCC) algorithm for acoustic feature extraction; continuous noise-level estimation to eliminate background noise; coarse- and fine-word boundary detection to define the word boundaries and Dynamic Time Warping algorithm to identify the words used.

That algorithm compares a series of energy vectors with unequal length and with duration variations within the series. It takes a weighted average difference between the feature vectors of the compared utterances and compares it with vectors in a template. The result is 97 to 100 percent accuracy for commands in the template...

...said Bloemandael. For example, echo cancellation with the SRS core would require only 1,000 additional gates. Complete OEM systems are also available that include

microphone, speaker,
battery and packaging.

October 25, 1999

13/3,K/2 (Item 1 from file: 16) Links
Gale Group PROMT(R)

06754707 Supplier Number: 56895049 (USE FORMAT 7 FOR FULLTEXT)

Frontier Design targets high-volume, command-and-control apps -- Speech-recognition core to open new markets.(demonstration at DSP World)(Company Business and Marketing)
Ohr, Stephan
Electronic Engineering Times , p 38
Oct 25 , 1999
Language: English Record Type: Fulltext
Document Type: Magazine/Journal ; Trade
Word Count: 990
-

...in a number of "giveaway" applications. "Talk to your postcard, and it will respond," said Mark Bloemandael, Frontier's applications manager in the Netherlands. More likely, it will find a place in speech-controlled car radios, an application that demands noise cancellation, high-accuracy speech recognition and low cost all in...

...a voice-controlled currency translation device for Columns Ltd. of Singapore. The design includes a complete speech recognition and synthesis (SRS) system on a chip, microphone and speaker. It includes 30 kbytes of RAM for the storage of speech recognition templates, and an additional 20 kbytes of ROM for the storage...
...other on-chip logic) or as a cell-based SoC including codec and amplifiers. Alternatively, Frontier can market a complete OEM module that includes speaker, microphone, IFR, RF and other functionality.

Since the speech recognition algorithm requires only 5 to 10 Mips, any existing pager, mobile telephone or other system with...

...of spare processing power can include the SRS core with no extra overhead.

The SRS implements several advanced recognition algorithms: the Mel Frequency Cepstrum Coefficient (MFCC) algorithm for acoustic feature extraction; continuous noise-level estimation to eliminate background noise; coarse- and fine-word boundary detection to define the word boundaries and Dynamic Time Warping algorithm to identify the words used.

That algorithm compares a series of energy vectors with unequal length and with duration variations within the series. It takes a weighted average difference between the feature vectors of the compared utterances and compares it with vectors in a template. The result is

97 to 100 percent accuracy for commands in the template...

...said Bloemandael. For example, echo cancellation with the SRS core would require only 1,000
additional gates. Complete OEM systems are also available that include microphone, speaker,
battery and packaging.

13/3,K/3 (Item 1 from file: 20) Links

07905898 (USE FORMAT 7 OR 9 FOR FULLTEXT)
Speech Recognition / Synthesis Core Delivers 97% to 100% -2-

BUSINESS WIRE
October 25, 1999
Journal Code: WBWE Language: English Record Type: FULLTEXT
Word Count: 806
(USE FORMAT 7 OR 9 FOR FULLTEXT)

-

Advanced Voice Recognition Algorithms -- Frontier Design employs the Mel Frequency Cepstrum
Coefficient (MFCC) algorithm for acoustic feature extraction, continuous noise level
estimation to eliminate background noise; coarse and fine word boundary detection to define
the word boundaries, and Dynamic Time Warping algorithm to identify the utterance.

-- Mel...

...the sensitivity of the human ear to frequency variations is equal across the spectrum. Mel
scaling results in less frequency sensitivity at high frequencies. The MFCC algorithm
consists of the calculation of an FFT power spectrum, followed by Mel scaling, log ii and an inverse
cosine transform (iDCT). This transformation is...

-- Continuous Noise Level Estimation -- The noise level estimation routine operates
continuously adapting to variations in the level of the background noise. It uses multiple
estimates and a selection algorithm to identify and eliminate background sounds and speech
artifacts (e.g. breathing, saying "uh").

-- Coarse Word Boundary Detection -- Coarse Word Boundary...

...duration characteristics of the audio signal

-- Fine Word Boundary Detection -- The fine Word Boundary Detection algorithm separates
irrelevant sounds (e.g. mouth clicks, breath noise, microphone rumble and background
sound)from the word by performing a detailed analysis of the energy levels

during and surrounding
the word.
-- Dynamic Time Warp Algorithm...

...identify the word. DTW compares of series of vectors with unequal length and with duration variations within the series. The resulting DTW distance is the weighted average difference between the feature vectors of the compared utterances, independent of their absolute time position in the energy pulse, but dependent of their relative position in the acoustical variations within...

...each in volumes of 1,000 units, plus $25,000 to $75,000 for design services, and $30,000 for ASIC prototypes. ASIC packaging adds approximately $0.40 per unit. Buyout license schemes are also offered.

Frontier Design was founded in 1997 to develop a next generation system level design methodology...

13/3,K/4 (Item 1 from file: 112) Links
UBM Industry News

01227574 (USE FORMAT 7 OR 9 FOR FULLTEXT)
Speech recognition/synthesis core has soft or hard options

Electronic Engineering , p 9
November, 1999
Language: English Record Type: Fulltext Document Type: Journal
Word Count: 00003201 (USE FORMAT 7 OR 9 FOR FULLTEXT)

Text:

...without interfaces to other on-chip logic; as a cell-based SOC including codec and amplifiers; or in a complete OEM module that includes speaker, microphone, IFR, RF and other functionality. The C-language core currently runs on DSP processors from Texas Instruments, the TI320C62XX, and National Semiconductor, CR16B DECT core...

...on-a-chip (SoC) implementations, along the lines of that shown in figure 1, that incorporate the speech recognition memory, synthesis ROM memory, AD and microphone amplifier and PWM speaker stage. The SoC requires no external components.

Additional functionality, such as echo cancellation, speech compression, or caller-id detection can be...

...count.

For example, adding echo cancellation with the SRS core would require only 1,000 additional gates. Complete OEM systems are also available that include microphone, speaker, battery and

packaging.
Frontier Design has developed one such speech recognition and synthesis system for Columns
Ltd. of Singapore.
Its first application is a voice-controlled currency translation device.
The design includes a complete speech recognition system-on-a-chip, microphone, and
speaker.
The speech recognition SoC was designed using Frontier's C-to-Silicon design methodology and
includes the SRS co-processor, ADC, DAC, serial...

...planned battery life of two years.
How it works
The Frontier Design SRS employs an advanced voice recognition algorithm, called the Mel
Frequency Cepstrum Coefficient (MFCC) algorithm, which is used for acoustic feature
extraction, continuous noise level estimation to eliminate background noise; coarse and fine
word boundary detection to define the word boundaries, and Dynamic Time Warping algorithm to
identify the utterance.
The MFCC algorithm uses the Mel scale which is a frequency scale in which the
sensitivity of the human ear to frequency variations is equal across the...

...it is a frequency scale in which the sensitivity of the human ear is equal across the spectrum.
As shown schematically in figure 3, the MFCC algorithm consists of the calculation of
an FFT power spectrum, followed by Mel scaling, log normalization and an inverse cosine transform
(iDCT).
This transformation is performed on overlapping frames of samples that have been Hamming
windowed.The default dimension of the MFCC data vectors (ie, feature vectors) is 16 elements
plus the logarithm of the signal energy of the frame.
The dimension of the data vector is...

...which is traded-off for degraded recognition results (a value above ten is recommended).
An example of the same spectrum as it would appear after cepstral averaging using 16
and 8 cepstral coefficients is shown in figure 4.
The noise level estimation routine operates continuously adapting to variations in the
level of the background noise. It uses multiple estimates and a selection algorithm to
identify and eliminate background sounds and speech artifacts (eg, breathing, saying "uh").
Coarse Word Boundary Detection determines when a whole Word Boundary Detection algorithm
separates irrelevant sounds (eg, mouth clicks, breath noise, microphone rumble and
background sound) from the word by performing a detailed analysis of the energy levels during and
surrounding the word.
The Dynamic Time Warp...

...identify the word. DTW compares a series of vectors with unequal length and with duration

variations within the series.
The resulting DTW distance is the weighted average difference between the feature vectors of the compared utterances, independent of their absolute time position in the energy pulse, but dependent of their relative position in the acoustical variations within...

...in volumes of 1,000 units, plus $25,000 to $75,000 for design services, and about $30,000 for ASIC prototypes. ASIC packaging adds approximately $0.40 per unit. Buyout license schemes are also offered.
Speech recognition the detail
The block schematic of the complete online processing with an expanded...

...of the speech.
In order to find the word boundaries, it is necessary to know the sound level of the background noise. The noise level estimation routine operates continuously, adapting to variations in the level of the background noise when necessary. The noise level estimation routine uses multiple estimates and a selection algorithm to use the best estimate for good performance in the presence of variations in the background noise, background sounds, speech and speech artefacts (such as breathing). The measured noise level is also used as an input to the AGC in order to minimize the effect of background noise level variations. The continuous noise level estimation works in cooperation with the coarse word boundary detection to classify between speech and non-speech states.
Word boundaries
With information about the background noise and its variations, it is possible to make a first rough estimate of the word boundaries. This part of the endpoint detection is performed on-line, which means that no samples can be lost during this part...

...example of coarse word boundary detection is shown in figure 5.
There are often spurious sounds surrounding a word, such as mouth clicks, breath noise, microphone rumble, and background sounds. The fine word boundary detections performs a detailed analysis of the energy levels during and surrounding a word in order to...techniques to enable comparison of a series of vectors with unequal length and with duration variations within the series. The resulting DTW distance is the weighted average difference between the feature vectors of the compared utterances, independent of their absolute time position in the energy pulse, but dependent of their relative position in the acoustical variations within...

...vocabulary number of the recognized word or an indication that the word has been rejected.
Quality analysis
An expression for the confidence interval of the probability of success p has been derived from the binomial distribution for p and application of the maximum

log-likelihood
principle, yielding that shown in equation 1.
N indicates the number of tests performed (ie, number of different template sets used), n is
the total number of utterances supplied to the core to be recognized, p is the actual
probability of successful recognition, pe is the estimated probability as found from the rate of success in the experiments, ua/2 is the right-critical value for a
reliability of the confidence interval of (1-a).
Performing extensive recognition tests an estimate of p (pe) can be obtained. From the
above expression a confidence interval for p can be found, so that the reliability of the
estimated probability is found. Using a critical value of 1.645 a reliability of 90%
is obtained.
For quality analysis for voice controlled dialling several tests have...

...on a set of 1197 words being on average 46 utterances of the 26 trained words. The language was
Dutch, the speaker was male, SNR approximately 35dB. Recordings were performed in an office
environment using a consumer quality microphone. The sample frequency was 8kHz, recorded in
office environment with low budget consumer quality microphone. The words trained were a
mixture of digits and names: "nul, een, twee, drie, vier, vijf, zes, zeven, acht, negen, piet,
klaas, gerard, anje, arend...

...mistake and 11 rejects. Therefore (p) = 99.0% +/-0.7% with a reliability of 90%. When correcting
the score for five rightfully rejects audio artefacts (microphone clicks, sighs, background
noise) the score is (p) 99.4 +/-0.7% with a reliability of 90%. The number of errors (mistaken
interpreted words) was...

13/3,K/5 (Item 1 from file: 647) Links
CMP Computer Fulltext

01203162 CMP Accession Number: EET19991025S0040
Frontier Design targets high-volume, command-and-control apps - Speech-recognition core to open new markets

Stephan Ohr
ELECTRONIC ENGINEERING TIMES , 1999 , n 1084 , PG38
Publication Date: 991025
Journal Code: EET Language: English
Record Type: Fulltext
Section Heading: International
Word Count: 995
Text:

...in a number of "giveaway" applications. "Talk to your postcard, and it will respond," said Mark
Bloemandael, Frontier's applications manager in the Netherlands. More likely, it will find a

place in speech-controlled car radios, an application that demands noise cancellation,
high-accuracy speech recognition and low cost all in...

...a voice-controlled currency translation device for Columns Ltd. of Singapore. The design
includes a complete speech recognition and synthesis (SRS) system on a chip, microphone and
speaker. It includes 30 kbytes of RAM for the storage of speech recognition templates, and an
additional 20 kbytes of ROM for the storage...

...other on-chip logic) or as a cell-based SoC including codec and amplifiers. Alternatively,
Frontier can market a complete OEM module that includes speaker, microphone, IFR, RF and
other functionality.

Since the speech recognition algorithm requires only 5 to 10 Mips, any existing pager,
mobile telephone or other system with...

...of spare processing power can include the SRS core with no extra overhead.

The SRS implements several advanced recognition algorithms: the Mel Frequency Cepstrum
Coefficient (MFCC) algorithm for acoustic feature extraction; continuous noise-level
estimation to eliminate background noise; coarse- and fine-word boundary detection to define
the word boundaries and Dynamic Time Warping algorithm to identify the words used.

That algorithm compares a series of energy vectors with unequal length and with duration
variations within the series. It takes a weighted average difference between the feature
vectors of the compared utterances and compares it with vectors in a template. The result
is 97 to 100 percent accuracy for commands in the template...

...said Bloemandael. For example, echo cancellation with the SRS core would require only 1,000
additional gates. Complete OEM systems are also available that include microphone , speaker,
battery and packaging.

19/3,K/1 (Item 1 from file: 88) Links
Gale Group Business A.R.T.S.

04894809 Supplier Number: 21094803
Reconstructions of truncated projections using an optimal basis expansion derived from the cross-correlation of a "knowledge set" of a priori cross-sections.(Selected Papers from the 1997 Medical Imaging Conference (MIC))
Panin, V.Y.; Zeng, G.L.; Gullberg, G.T.
IEEE Transactions on Nuclear Science , v45 , n4 , p2119(7)
August , 1998
ISSN: 0018-9499
Language: English Record Type: Abstract
...optimal basis expansion derived from the cross-correlation of a "knowledge set" of a priori cross-sections.(Selected Papers from the 1997 Medical Imaging Conference (MIC)) (
Author Abstract: ...the coefficients of this expansion by minimizing the sum of squares difference between the expansion and the projection measurements taking into account the distribution of coefficients over basis vectors. The

constrained least-squares estimates of the coefficients were used in an expansion of the orthogonal basis to obtain the reconstructed image. The constrained solution has a reduced noise level in this inverse problem. It is shown that the reconstruction of truncated projections can be significantly improved over that of commonly used iterative reconstruction..